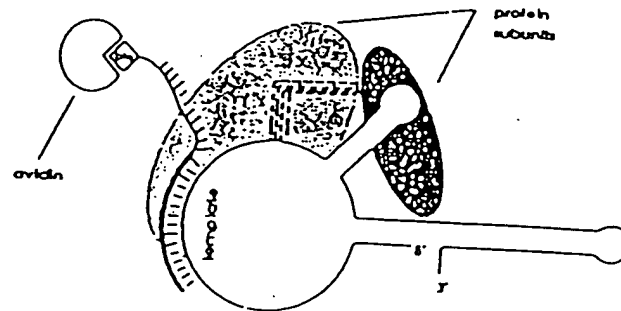


FIGURE 1

PANEL A



PANEL B

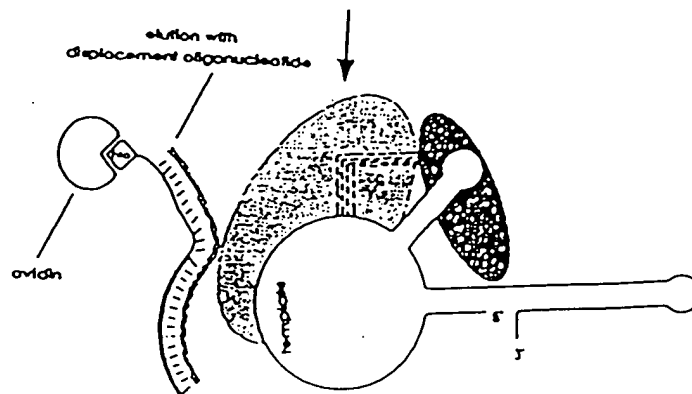


FIGURE 2

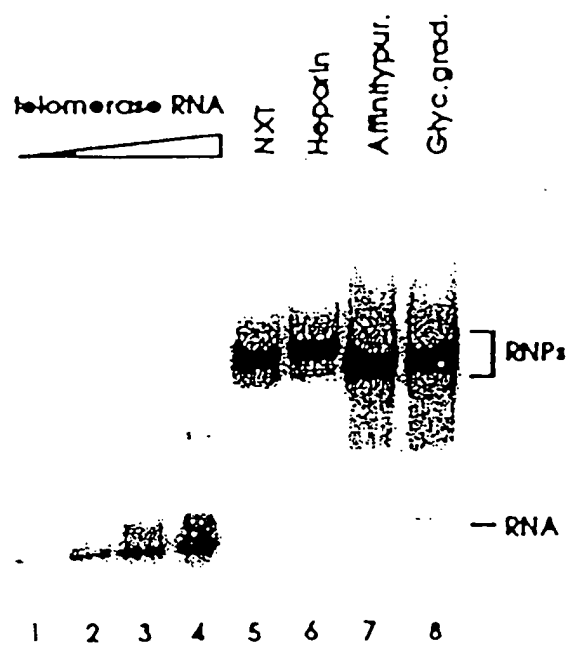


FIGURE 3

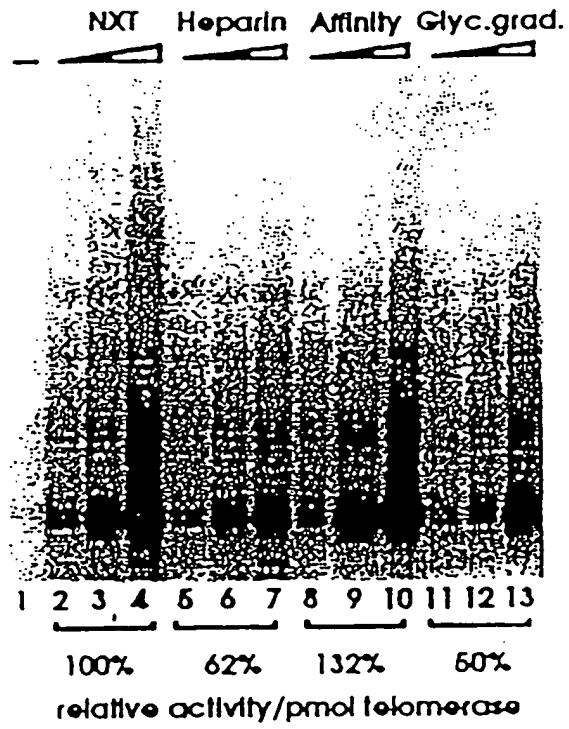


FIGURE 4

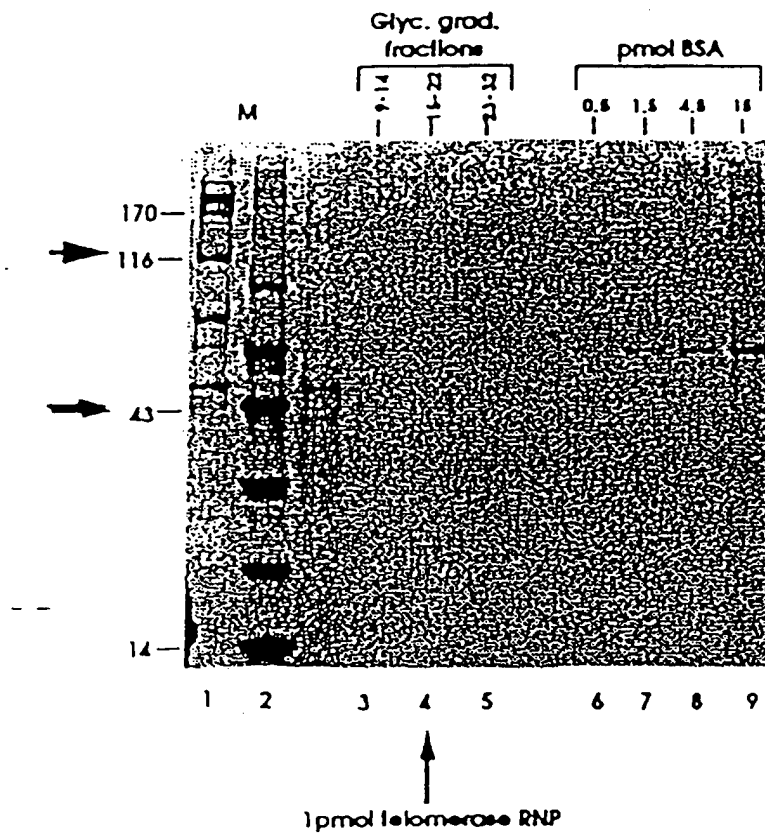


FIGURE 5

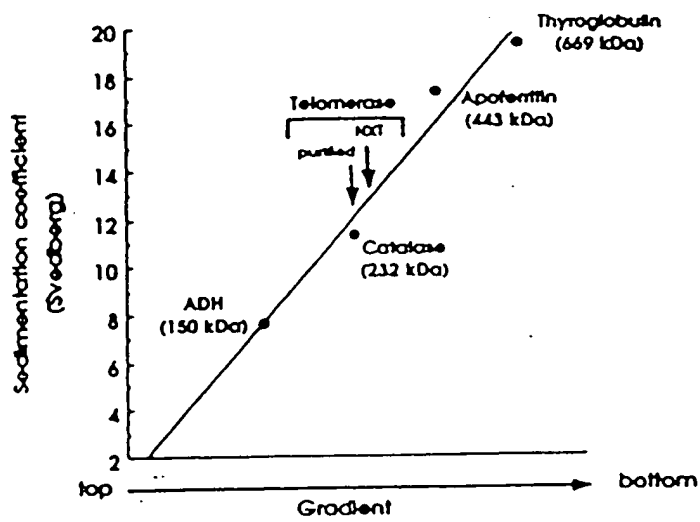


FIGURE 6

Photo: 9292700

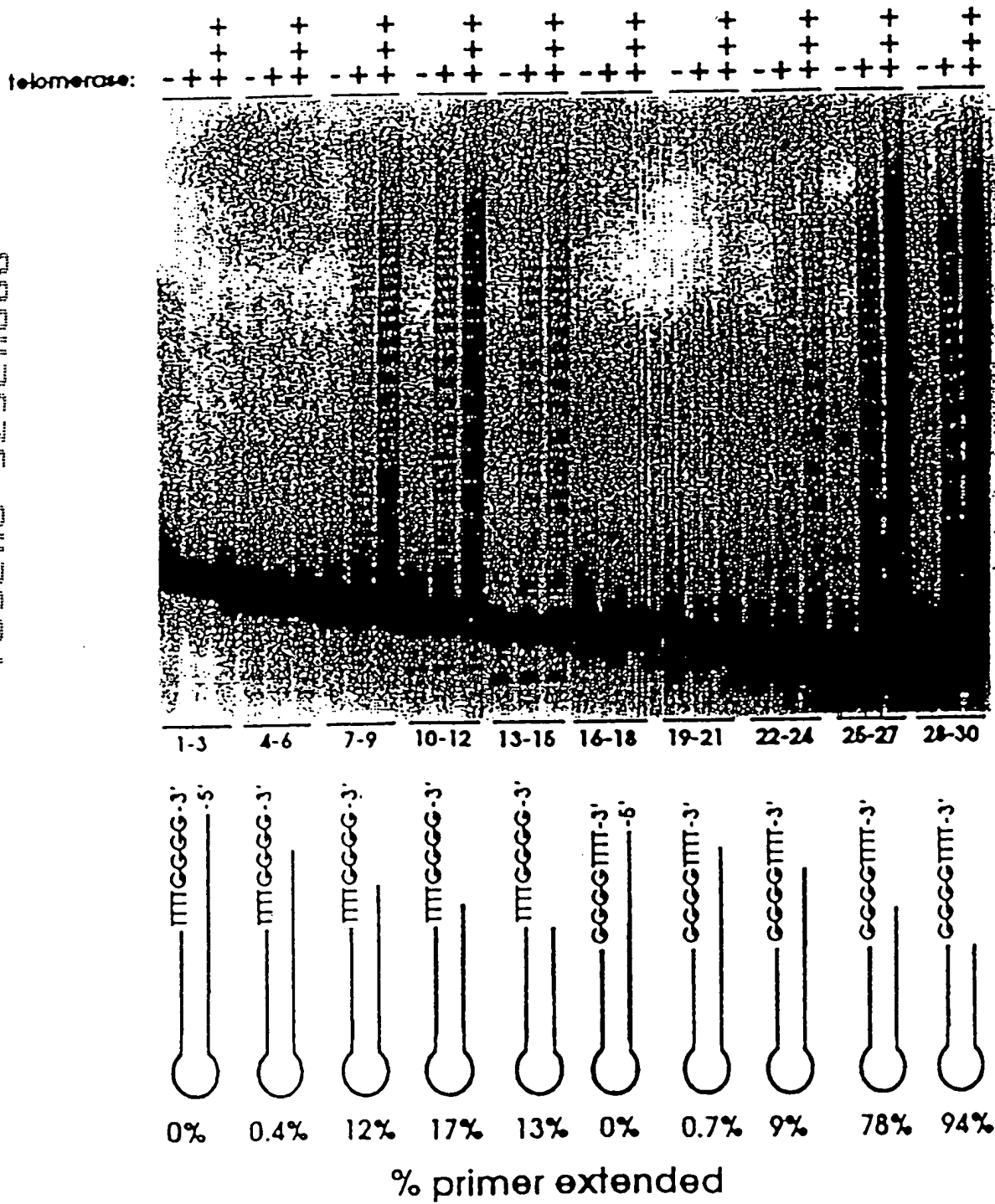


FIGURE 7

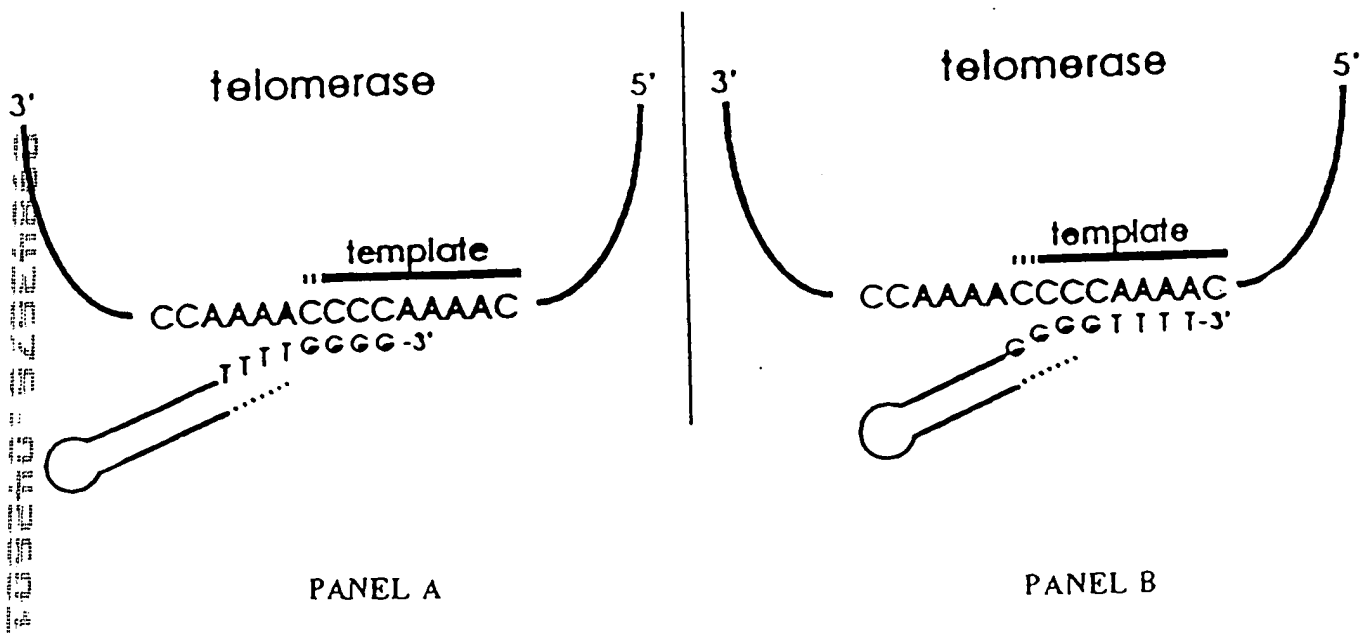


FIGURE 8

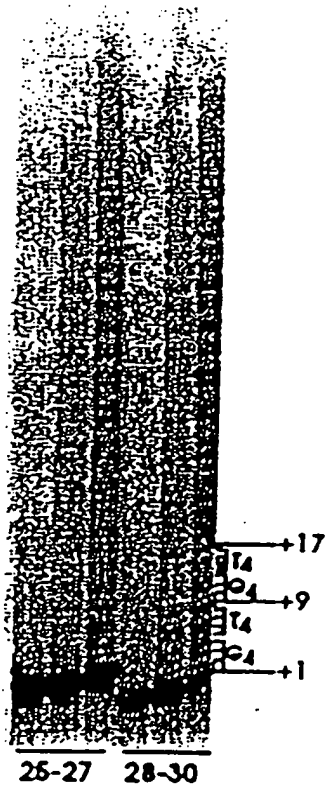


FIGURE 8

FIGURE 9

1 AAAACCCCAA AACCCCAAAA CCCCTTTTAG AGCCCTGCAG TTGGAAATAT
 51 AACCTCAGTA TTAATAAGCT CAGATTTTAA ATATTAATTA CAAAACCTAA
 101 ATGGAGGTTG ATGTTGATAA TCAAGCTGAT AATCATGGCA TTCACTCAGC
 151 TCTTAAGACT TGTGAAGAAA TTAAAGAAGC TAAAACGTTG TACTCTTGGA
 201 TCCAGAAAGT TATTAGATGA AGAAATCAAT CTCAAAGTCA TTATAAAGAT
 251 TTAGAAGATA TTAAAATATT TGC GCAGACA AATATTGTTG CTACTCCACG
 301 AGACTATAAT GAAGAAGATT TTAAAGTTAT TGCAAGAAAA GAAGTATTTT
 351 CAACTGGACT AATGATCGAA CTTATTGACA AATGCTTAGT TGAACCTTCT
 401 TCATCAAGCG ATGTTTCAGA TAGACAAAAA CTTCAATGAT TTGGATTTC
 451 ACTTAAGGGA AATCAATTAG CAAAGACCCA TTTATTAACA GCTCTTTCAA
 501 CTCAAAGCA GTATTTCTTT CAAGACGAAT GGAACCAAGT TAGAGCAATG
 551 ATTGGAAATG AGCTCTTCCG ACATCTCTAC ACTAAATATT TAATATTCCA
 601 GCGAACTTCT GAAGGAACTC TTGTTCAATT TTGCGGGAAT AACGTTTTTG
 651 ATCATTTGAA AGTCAACGAT AAGTTTGACA AAAAGCAAAA AGGTGGAGCA
 701 GCAGACATGA ATGAACCTCG ATGTTGATCA ACCTGCAAAT ACAATGTCAA
 751 GAATGAGAAA GATCACTTTC TCAACAACAT CAACGTGCCG AATTGGAATA
 801 ATATGAAATC AAGAACCAGA ATATTTTATT GCACTCATTT TAATAGAAAT
 851 AACCAATTCT TCAAAAAGCA TGAGTTTG TG AGTAACAAAA ACAATATTTT
 901 AGCGATGGAC AGAGCTCAGA CGATATTCAC GAATATATTC AGATTTAATA
 951 GAATTAGAAA GAAGCTAAAA GATAAGGTTA TCGAAAAAAT TGCCTACATG
 1001 CTTGAGAAAG TCAAAGATTT TAACTTCAAC TACTATTTAA CAAAATCTTG
 1051 TCCTCTTCCA GAAAATTGGC GGAACGGAA AAAAAAATC GAAAACCTGA
 1101 TAAATAAAAC TAGAGAAGAA AAGTCGAAGT ACTATGAAGA GCTGTTTAGC
 1151 TACACAACCTG ATAATAAATG CGTCACACAA TTTATTAATG AATTTTTCTA
 1201 CAATATACTC CCCAAAGACT TTTTGACTGG AAGAAACCGT AAGAATTTTC
 1251 AAAAGAAAGT TAAGAAATAT GTGGAACATA ACAAGCATGA ACTCATTAC
 1301 AAAAACTTAT TGCTTGAGAA GATCAATACA AGAGAAATAT CATGGATGCA
 1351 GGTTGAGACC TCTGCAAAGC ATTTTATTAT TTTTGATCAC GAAAACATCT
 1401 ACGTCTTATG GAAATTGCTC CGATGGATAT TCGAGGATCT CGTCGTCTCG
 1451 CTGATTAGAT GATTTTTCTA TGTCACCGAG CAACAGAAAA GTTACTCCAA
 1501 AACCTATTAC TACAGAAAGA ATATTTGGGA CGTCATTATG AAAATGTCAA
 1551 TCGCAGACTT AAAGAAGGAA ACGCTTGCTG AGGTCCAAGA AAAAGAGGTT
 1601 GAAGAATGGA AAAAGTCGCT TGGATTGCA CCTGGAAAAC TCAGACTAAT
 1651 ACCGAAGAAA ACTACTTTCC GTCCAATTAT GACTTTCAAT AAGAAGATTG
 1701 TAAATTCAGA CCGGAAGACT ACAAATTA CTACAAATAC GAAGTTATTG
 1751 AACTCTCACT TAATGCTTAA GACATTGAAG AATAGAATGT TTAAAGATCC
 1801 TTTTGGATTG GCTGTTTTTA ACTATGATGA TGTAATGAAA AAGTATGAGG
 1851 AGTTTGTGTTG CAAATGGAAG CAAGTTGGAC AACCAAACT CTTCTTTGCA
 1901 ACTATGGATA TCGAAAAGTG ATATGATAGT GTAAACAGAG AAAAATATC
 1951 AACATTCCTA AAAACTACTA AATTACTTTC TTCAGATTTT TGGATTATGA
 2001 CTGCACAAAT TCTAAAGAGA AAGAATAACA TAGTTATCGA TTCGAAAAAC
 2051 TTTAGAAAGA AAGAAATGAA AGATTATTTT AGACAGAAAT TCCAGAAGAT
 2101 TGCATTGAA GGAGGACAAT ATCCAACCTT ATTCAGTGTT CTTGAAAATG
 2151 AACAAAATGA CTTAAATGCA AAGAAAACAT TAATTGTTGA AGCAAAGCAA
 2201 AGAAATTATT TTAAGAAAGA TAACTTACTT CAACCAGTCA TTAATATTTG
 2251 CCAATATAAT TACATTAAT TTAATGGGAA GTTTTATAAA CAAACAAAAG
 2301 GAATTCCTCA AGGTCTTTGA GTTTCATCAA TTTTGTATC ATTTTATTAT

[illegible]

2351 GCAACATTAG AGGAAAGCTC CTTAGGATTC CTTAGAGATG AATCAATGAA
2401 CCCTGAAAAT CCAAATGTTA ATCTTCTAAT GAGACTTACA GATGACTATC
2451 TTTTGATTAC AACTCAAGAG AATAATGCAG TATTGTTTAT TGAGAAACTT
2501 ATAAACGTAA GTCGTGAAAA TGGATTTAAA TTCAATATGA AGAAACTACA
2551 GACTAGTTTT CCATTAAGTC CAAGCAAATT TGCAAAATAC GGAATGGATA
2601 GTGTTGAGGA GCAAAATATT GTTCAAGATT ACTGCGATTG GATTGGCATC
2651 TCAATTGATA TGAAAACCTCT TGCTTTAATG CCAAATATTA ACTTGAGAAT
2701 AGAAGGAATT CTGTGTACAC TCAATCTAAA CATGCAAACA AAGAAAGCAT
2751 CAATGTGGCT CAAGAAGAAA CTAAAGTCGT TTTTAATGAA TAACATTACC
2801 CATTATTTTA GAAAGACGAT TACAACCGAA GACTTTGCGA ATAAAACCTCT
2851 CAACAAGTTA TTTATATCAG GCGGTTACAA ATACATGCAA TGAGCCAAAG
2901 AATACAAGGA CCACTTTAAG AAGAAGCTTAG CTATGAGCAG TATGATCGAC
2951 TTAGAGGTAT CTAAAATTAT ATACTCTGTA ACCAGAGCAT TCTTTAAATA
3001 CCTTGTGTGC AATATTAAGG ATACAATTTT TGGAGAGGAG CATTATCCAG
3051 ACTTTTTCTT TAGCACACTG AAGCACTTTA TTGAAATATT CAGCACAAAA
3101 AAGTACATTT TCAACAGAGT TTGCATGATC CTCAAGGCAA AAGAAGCAAA
3151 GCTAAAAAGT GACCAATGTC AATCTCTAAT TCAATATGAT GCATAGTCGA
3201 CTATTCTAAC TTATTTTGGG AAGTTAATTT TCAATTTTTG TCTTATATAC
3251 TGGGGTTTTG GGGTTTTGGG GTTTTGGGG

FIGURE 10

1 MEVDVDNQAD NHGIHSALKT CEEIKEAKTL YSWIQKVIRC RNQSQSHYKD
 51 LEDIKIFAQT NIVATPRDYN EEDFKVIARK EVFSTGLMIE LIDKCLVELL
 101 SSSDVSDRQK LQCFGFQLKG NQLAKTHLLT ALSTQKQYFF QDEWNQVRAM
 151 IGNELFRHLY TKYLIFORTS EGTLVQFCGN NVFDHLKVND KFDKKQKGGA
 201 ADMNEPRCCS TCKYNVKNK DHFLNNINVP NWNNMKSRTTR IFYCTHFNRN
 251 NQFFKKHEFV SNKNNISAMD RAQTIFTNIF RFNRIRKKLK DKVIEKJAYM
 301 LEKVKDFNFN YYLTKSCPLP ENWRERKQKI ENLINKTREE KSKYYEELFS
 351 YTTDNKCVTQ FINEFFYNIL PKDFLTGRNR KNFQKKVKKY VELNKHELH
 401 KNLLEKINT REISWMQVET SAKHFYFDH ENIYVLWKL RWFEDLVVS
 451 LIRCFYVTE QQKSYSKTY YRKNWDVIM KMSIADLKKE TLAEVQEKEV
 501 EEWKKS LGFA PGKRLIPKK TFRPIMTFN KKIVNSDRKT TKLTTNTKLL
 551 NSHMLMLTKL NRMFKDPFGF AVFNYDDVMK KYEEFVCKWK QVGQPKLFFA
 601 TMDIEKCYDS VNREKLSTFL KTTKLLSSDF WIMTAQILKR KNNIVIDSKN
 651 FRKKEMKD YF RQKFQKIALE GGQYPTLFSV LENEQNDLNA KKTLLIVEAKQ
 701 RNYFKKDNLL QPVINICQYN YINFNGKFYK QTKGIPQGLC VSSILSSFY
 751 ATLEESSLGF LRDESMNPEN PNVNLLMRLT DDYLLITTQE NNAVLFIEKL
 801 INVSRENGFK FNMKKLQTSF PLSPSKFAKY GMDSVEEQNI VQDYCDWIGI
 851 SIDMKTLALM PNINLRIEGL CTLNLNMQT KKASMWLKKK LKSFLMNNIT
 901 HYFRKTTITTE DFANKTLNKL FISGGYKYMQ CAKEYKDHFK KNLAMSSMID
 951 LEVSKIIYSV TRAFFKYLVC NIKDTIFGEE HYPDFFLSTL KHFIIEFSTK
 1001 KYIFNRVCM I LKAKEAKLKS DQCQSLIQYD A

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 12

CCCCCAACCCCAACCCCAACCCCTATAAAAAAGAAAAATTGAGGTAGTTTAGA
 1 ----- 60
 GGGGTTTGGGGTTTGGGGTTTGGGGATATTTTTCTTTTAACTCCATCAAATCT
 a P Q N P K T P K P L * K K K K L R * F R -
 b P K T P K P Q N P Y K K R K N C G S L E -
 c P K P Q N P K T P I K K E K I E V V * K -
 AATAAAATATTATCCCGCACAAATGGAGATGGATATTGATTGGATGATATAGAAAATT
 61 ----- 120
 TTATTTTATAATAAGGCGTGTACCTCTACCTATAACTAACTACTATATCTTTTAA
 a N K I L F P H K W R W I L I W H I * K I -
 b I K Y Y S R T N G D G Y C F G C Y R K F -
 c * N I I P A Q M E M D I D L D D I E N L -
 TACTTCTAATACATTCAACAAGTATAGCAGCTCTTGTAGTGACAAGAAAGGATGCAAAA
 121 ----- 180
 ATGAAGGATTATGTAAGTTGTCATATCGTCGAGAACATCACTGTTCTTCTACGTTTT
 a Y F L I H S T S I A A L V V T R K D A K -
 b T S * Y I Q Q V * Q L L * Q E R M Q N -
 c L P N T F N K Y S S S C S D K K G C K T -
 CATTGAAATCTGGCTCGAAATCGCCTTCATTGACTATTCCAAAGTTGCAAAAACAATTAG
 181 ----- 240
 GTAACTTTAGACCGAGCTTTAGCGGAAGTAACTGATAAGGTTTCAACGTTTTTGTAAATC
 a H C N L A R N R L H C L F Q S C K N N * -
 b I E I W L E I A F I D Y S K V A K T I R -
 c L K S G S K S P S L T I P K L Q K Q L E -
 AGTTCTACTTCTCGGATGCAAACTTTTATAACGATTCTTCTTGAGAAAATTAGTTTTAA
 241 ----- 300
 TCAAGATGAAGAGCCTACGTTTAGAAATATTGCTAAGAAAGAACTCTTTTAATCAAAATT
 a S S T S R H Q I F I T I L S C E N * F * -
 b V L L L G C X S L * R F F L E K I S F K -
 c P Y F S D A N L Y N D S F L R K L V L K -
 AAAGCGAGAGCAAGAGTAGAAATGAAACATTACTAATGTTTAAATAAAATCAGGTAA
 301 ----- 360
 TTTCCGCTCTCGTTTCTCATCTTTAACTTTGTAATGATTACAAATTTATTTAGTCCATT
 a K A E S K E * K L K H Y * C L N K I R * -
 b K R R A K S R N C N I T N V * I K S G N -
 c S G E Q R V E I E T L L M F K * N Q V H -
 TGAGGATTATTCTATTTTTTAGATCACTTCTTAAGGAGCATTATGGAGAAAATTACTTAA
 361 ----- 420
 ACTCCTAATAAGATAAAAAATCTAGTGAAGAATTCCTCGTAATACCTCTTTTAATGAATT
 a C G L F Y F L D H F L R S I H E K I T * -
 b E D Y S I F * I T S * G A L W R K L L N -
 c R I I L F F R S L L K E H Y G E N Y L I -
 TACTAAAAGTAAACAGTTTGGATTATTTCCCTAGCCAAACATGATGAGTATATTAAATT
 421 ----- 480
 ATGATTTTCCATTGTGCAACCTAATAAAGGGATCGGTTGTTACTACTCATATAATTAA
 a Y * K V N S L D Y F P S Q Q C C V Y * I -
 b T K R * T V W I I S L A N N D E Y I K F -
 c L K G K Q F G L F P * P T M M S I L N S -

FIGURE 12 (cont.)

CATATGAGAATGAGTCAAAGGATCTCGATACATCAGACTTACCAAAGACAACTCGCTAT
 481 ----- 540
 GTATACTCTTACTCAGTTTCCTAGAGCTATGTAGTCTGAATGGTTCTGTTTGAGCGATA

 a H M R M S Q R I S I H O T Y Q R Q T R Y -
 b I C E C V K G S R Y I R L T K D K L A I -
 c Y E N E S K D L D T S D L P K T N S L -

 AAAACGCAAGAAAAAGTTTGATAATCGAACAGCAGAAGAACTTATTGCATTACTATTTCG
 541 ----- 600
 TTTTGGCTTCTTTTCAAACCTATTAGCTTGTCTGCTTCTTGAATAACGTAATGATAAGC

 a K T Q E K V C * S N S R R T Y C I Y Y S -
 b K R K K K F D N R T A E E L I A F T I R -
 c N A R K S L I I E Q Q K N L L H L L F V -

 TATGGGTTTTATTACAATTGTTTTAGGTATCGACGGTGAACCTCCGAGTCTTGAGACAAT
 601 ----- 660
 ATACCCAAAATAATGTTAACAATAATCCATAGCTGCCACTTGAGGGCTCAGAAGCTCTGTTA

 a Y G F Y Y N C F R Y R R C T P E S C D N -
 b M G F I T I V L G I D G E L P S L E T I -
 c W V L L Q L F * V S T V N S R V L R Q L -

 TGAAAAAGCTGTTTACAACCTGAAGGAATCGCAGTTCTGAAAGTTCTGATGTGTATGCCAT
 661 ----- 720
 ACTTTTTTCACAATGTTGACTTCCTTAGCGTCAAGACTTTCAAGACTACACATACGGTA

 a C K S C L Q L K E S Q F C K F * C V C H -
 b E K A V Y N C R N R S S E S S D V Y A I -
 c K K L F T T E G I A V L K V L M C M P L -

 TATTTTGTAATTAATCTCAAATATCTTATCTCAATTTAATGGATAGCTATAGAAACAAA
 721 ----- 780
 ATAAACACTTAATTAGAGTTTATAGAATAGAGTTAAATTACCTATCGATATCTTTGTTT

 a Y F V N * S Q I S Y L N L M D S Y R N K -
 b I L C I N L K Y L I S I * W I A I E T N -
 c F C E L I S N I L S Q F N G * L * K Q T -

 CCAAATAAACCATGCAAGTTTAATGGAATATACGTTAAATCCTTTGGGACAAATGCACAC
 781 ----- 840
 GGTTTATTTGGTACGTTCAAATTACCTTATATGCAATTTAGGAAACCTGTTTACGTGTG

 a P N K P C K F N G I Y V K S F G T N A H -
 b Q I N H A S L M E Y T L N P L G Q M H T -
 c K * T M Q V * W N I R * I L W D K C T L -

 TGAATTTTATTTGGATTCTTAAAGCATAGATACAGAAATGCTTTAGAGACTGATTTAGC
 841 ----- 900
 ACTTAAATATAACCTAAGAAATTCGTATCTATGTCTTACGAAATCTCTGACTAAATCG

 a C I Y I G F L K H R Y T E C F R D C F S -
 b E F I L D S * S I D T Q N A L E T D L A -
 c N L Y W I L R A * I H R M L * R L I * L -

 TTACAACAGATTACCTGTTTTGATTACTCTTGCTCATCTCTTATATCTTTAAAAGAAGCA
 901 ----- 960
 AATGTTGCTCTAATGGACAAAACCTAATGAGAACGAGTAGAGAATATAGAAATTTCTTCGT

 a L Q Q I F C F D Y S C S S L I S L K E A -
 b Y N R L P V L I T L A H L L Y L * K K Q -
 c T T D Y L F C L L L I S Y I F K R S R -

 GGCGAAATGAAAAGAAGACTAAAGAAAAGAGATTTCAAAATTTGTTGATTCTTCTGTAACC
 961 ----- 1020
 CCGCTTTACTTTTCTTGATTCTTTCTCTAAAGTTTAAACAACTAAGAAGACATTGG

 a G E M K R R R L K K E I S K F V D S S V T -
 b A K C K E D * R K R F Q N L L I L L * P -
 c R N E K K T K E R D F K I C C F F C N R -

 GGAATTAACAACAAGAAATATTAGCAACGAAAAAGAAGAGCTATCACAATCCTGATTG
 1021 ----- 1080
 CCTTAATTGTTGTTCTTATAATCGTTGCTTTTTCTTCTCTCGATAGTGTAGGACTAAG

 a G I N N K N I S N E K E E E L S O S C F -
 b E L T T R I L A T K K K K S Y H N P D S -
 c N * Q Q E Y * Q R K R R R A I T I L I L -

1000
 900
 800
 700
 600
 500
 400
 300
 200
 100
 0

FIGURE 12 (cont.)

1081 TTAAGATTTCAAAAATTCAGGTAAGAGAGATACATTCATTAATAATTCATATATTATAG 1140
 AATTTCTAAAGTTTTTAAGGTCCATTCTCTCTATGTAAGTAATTTTAAGTATATAATATC

a L K I S K I P G K R D T F I K I H I L -
 b - R F Q K F Q V R E I H S L K F I Y Y S -
 c K D F K N S R - E R Y I H - N S Y I I V -

1141 TTTTTCATTTCACAGCTGTTATTTTCTTTTATCTTAACAATATTTTGTATTAGCTGGAA 1200
 AAAAAGTAAAGTGTCGACAATAAAAGAAAATAGAATTGTTATAAAAAACTAATCGACCTT

a F F I S Q L L F S F I L T I F F D - L E -
 b F S F H S C Y F L L S - Q Y F L I S W K -
 c F H F T A V I F F Y L N N I F C L A G S -

1201 GTAAAAAGTATCAAATAAGAGAAGCGCTAGACTGAGGTAACCTAGCTTATTCACATTCAT 1260
 CATTTTTCATAGTTTATCTCTTCGCGATCTGACTCCATTGAATCGAATAAGTGTAAGTA

a V K S I K - E K R - T E V T - L I H I H -
 b - K V S N K R S A R L R - L S L F T F I -
 c K K Y Q I R E A L D C G N L A Y S H S -

1261 AGATCGACCTTCATATATCCAATACGATGATAAGGAAACAGCAGTCATCCGTTTTAAAAA 1320
 TCTAGCTGGAAGTATATAGGTTATGCTACTATTCTTTGTCGTCAGTAGGCAAAATTTT

a R S T F I Y P I R C - G N S S H P F - K -
 b D R P S Y I Q Y D D K E T A V I R F K N -
 c I D L H I S N T M I R K Q Q S S V L K I -

1321 TAGTGCTATGAGGACTAAATTTTATAGAGTCAAGAAATGGAGCCGAAATCTTAATCAAAAA 1380
 ATCAGGATACTCTGATTAAAAATCTCAGTTCTTTACCTCGGCTTAGAATTAGTTTTT

a - C Y E D - I F R V K K W S R N L N Q K -
 b S A H R T K F L E S R N G A E I L I K K -
 c V L C G L N F - S Q E M E P K S - S K R -

1381 GAATTGCGTCGATATTGCAAAAAGAATCGAACTCTAAATCTTTCGTTAATAAGTATTACCA 1440
 CTTAAGCGACGTATAACGTTTTCTTAGCTTGAGATTTAGAAAGCAATTATTCATAATGGT

a E L R R Y C K R I E L - I F R - - V L P -
 b N C V D I A K E S N S K S F V N K Y Y Q -
 c I A S I L Q K N R T L N L S L I S I T N -

1441 ATCTTGATTGATTGAAGAGATTGACGAGGCAACTGCACAGAAGATCATTAAAGAAATAAA 1500
 TAGAACTAACTAACTTCTCTAATCTGCTCGTTCGCTGCTTCTTAGTAATTCTTTATTT

a I L I D C R D - R G N C T E D H - R N K -
 b S C L I E E I D E A T A Q K I I K E I K -
 c L D C L K R L T R Q L H R R S L K K - S -

1501 GTAACCTTTTATTAATTAGAGAATAAACTAAATTACTAATATAGAGATCAGCGATCTTCAA 1560
 CATTCGAAATAATTAATCTCTTATTTGATTAAATGATTATATCTCTAGTCGCTAGAAAGT

a V T F I N - R I N - I T N I E I S D L Q -
 b - L L L I R E - T K L L I - R S A I F N -
 c N F Y - L E N K L N Y - Y R D Q R S S I -

1561 TTGACGAAATAAAAGCTGAACTAAAGTTAGACAATAAAAAATACAAACCTTGGTCAAAAT 1620
 AACTGCTTTATTTTCGACTTGATTTCATCTGTTATTTTATGTTTGGAAACCAGTTTAA

a L T K - K L N - S - T I K N T N L C Q N -
 b C R N K S C T K V R Q - K I O T L V K I -
 c D E I K A E L K L D N K K Y K P W S K Y -

1621 ATTGAGGAAGGAAAAGAAGACCAGTTAGCAAAAAGAAAAATAAGCGCAATAAATAAATGA 1680
 TAACTCCTTCTCTTCTCTGCTCAATCGTTTTCTTTTTTATTCGTTATTTATTTTACT

a I E E G K E D Q L A K E K I R O - I K C -
 b L R K E K K T S - Q K K K - G N K - N E -
 c C G R K R R P V S K R K N K A I N K H S -

8

```

1681      GTACAGAACTGAAGAAATAAAAGATTATTTTTCATAATTTATTGAAAAGAGGGGTT
-----> 1740
      CATGCTTCACTTCTTTATTTTCTAAATAAAAAAGTTATTAATAACTTTTCTCCCAA

a      V   Q   K   C   R   N   K   R   F   I   F   F   N   N   L   L   K   R   G   V   -
b      Y   R   S   E   E   I   K   D   L   F   F   S   I   I   Y   C   K   E   G   F   -
c      T   E   V   K   K   -   K   I   Y   F   F   Q   -   F   Y   I   E   K   R   G   F   -

      TTGGGGTTTTGGGGTTTTGGGG
1741 ----- 1762
      AACCCCAAAACCCCAAAACCCC

a      L   G   F   W   G   F   G   -
b      W   G   F   G   V   L   G   -
c      G   V   L   G   F   W   -

```


NAME	AGE	SEX	REL	DATE	TIME	PLACE	REMARKS
JOHN	25	M	H	1945	10:30	ST. LOUIS	ARRIVED
MARY	22	F	W	1945	11:00	ST. LOUIS	ARRIVED
JOHN	25	M	H	1945	11:30	ST. LOUIS	ARRIVED
MARY	22	F	W	1945	12:00	ST. LOUIS	ARRIVED
JOHN	25	M	H	1945	12:30	ST. LOUIS	ARRIVED
MARY	22	F	W	1945	13:00	ST. LOUIS	ARRIVED
JOHN	25	M	H	1945	13:30	ST. LOUIS	ARRIVED
MARY	22	F	W	1945	14:00	ST. LOUIS	ARRIVED
JOHN	25	M	H	1945	14:30	ST. LOUIS	ARRIVED
MARY	22	F	W	1945	15:00	ST. LOUIS	ARRIVED
JOHN	25	M	H	1945	15:30	ST. LOUIS	ARRIVED
MARY	22	F	W	1945	16:00	ST. LOUIS	ARRIVED
JOHN	25	M	H	1945	16:30	ST. LOUIS	ARRIVED
MARY	22	F	W	1945	17:00	ST. LOUIS	ARRIVED
JOHN	25	M	H	1945	17:30	ST. LOUIS	ARRIVED
MARY	22	F	W	1945	18:00	ST. LOUIS	ARRIVED
JOHN	25	M	H	1945	18:30	ST. LOUIS	ARRIVED
MARY	22	F	W	1945	19:00	ST. LOUIS	ARRIVED
JOHN	25	M	H	1945	19:30	ST. LOUIS	ARRIVED
MARY	22	F	W	1945	20:00	ST. LOUIS	ARRIVED
JOHN	25	M	H	1945	20:30	ST. LOUIS	ARRIVED
MARY	22	F	W	1945	21:00	ST. LOUIS	ARRIVED
JOHN	25	M	H	1945	21:30	ST. LOUIS	ARRIVED
MARY	22	F	W	1945	22:00	ST. LOUIS	ARRIVED
JOHN	25	M	H	1945	22:30	ST. LOUIS	ARRIVED
MARY	22	F	W	1945	23:00	ST. LOUIS	ARRIVED
JOHN	25	M	H	1945	23:30	ST. LOUIS	ARRIVED
MARY	22	F	W	1945	24:00	ST. LOUIS	ARRIVED

2	EVDVNDHADNHGHSALKTCEEIKEAKTLYSWIKVIRCNQOSQSHYKDL	51
19	ELELENOENQNDIOVRVK...IDDPKQY...LVNVTAACLLQEGSSYODK	62
52	EDIKIFAQTNIVATPRDYNEEDFKVIARKEVF..STGLMIELIDKCLVELL	100
63	DERRVITKALL...EVAESDPEFICQLAVYIRNELYIRTTTNYIVAF..	107
101	SSSDSDOROKLOCFGFQFKGNQLAKTHLLTALSTOKQYFFQDEWNQVRAM	150
108	CVVHKNTOPFIEKYFNKAVLLPNOLLEVECEFAQVLYI	144
151	IGNELFRHLYTKYLIFORTSEGLTVQFCGMNVFDHLKVNDKFKKQKQGA	200
145	FDATFKNLY...LDRILSQDIRKELTFRKCLQRCVRSKF	181
201	ADMNE...PRCCSTCKYNVKNKEDHFLNNINVPWNWNMKSRTIRIFYCTHF	247
182	SEFNEYQLGKYCTES...QRKKTFRYLSVTNKQKQWQTKKK	220
248	NRNQOFFKKHEFVSNKNNISAMDRAQITFTNIFRFRNIRKKLKDKVIEKI	297
221	RKENLLTKLOAIKESDKSKRETG...DIMNVEDAIKALPAVMKKI	264
298	AYHLEKVKDFNFNYLLTKSCPLPENWRERKOKIENLINKTREESKYEE	347
265	AKRQNAHK...KHMKAPKIPNSTLESKYLTFKD	294
348	LFSYTDNKCVTQFINEFFYNILPKDFLTGRNRKNFQKKVKKYVELNKHE	397
295	LIKFCHESEP...KERVYKILGKKYPKTEEEYKAAFSDSASAPFN..PE	338
398	LIHKNLLEKINTREISWMQVETSAKHFFYFDHENIYVLWKLRLWIFEDL	447
339	LAKRMIKIEISKWENELSAKGNTAEVWDNLISSNQLPYHMLRLNSL..	386
448	VVSLIRCFYVTEQOKSYSKTYYYRKNIWDVIMKMSIADLKKETLAEVQE	497
387	ILKAGVSD	394
498	KEVEEWKKSGLGFAFGKRLRLPKKTTFRPIMTFNKKIVNSDRKTKLTNT	547
395	TTHS	398
548	KLLNSHMLKTLKNRMFKDPFGFAVFNYDDVMKKYEEFVCKWKQVQPKL	597
399	IVINK...ICEPKAVENSKM	415
598	FFATMDIEKCYDSVNREKLSTFLKTTKLLSSDFWIMTAQILKRKNIVID	647
416	F..PLOFFSAIEAVN..EAVTKGFKAKK...RENMLKGQIEAVKE..VVE	457
648	SKNFRKKEMKDYFRQKFQKIALEGGQYPTLFSVLENEQNDLNAKKT LIVE	697
458	KTDEEKKDM...ELEQTEEGEFVKVNEGIGQYINSIELAIK	496
698	AKQRNYFKKDNLLQPVINICQYNYINFNGKPYKQTKGIPQGLCVSSILSS	747
497	IAVKNLDEIKGHTAIFSDVSGSMSTSMGGAKKYGSVRTCLECALVLGL	546
748	FYYATLEESSLGLRDESMNPENPNVNLMLRLTDDYLLITTOENNAVLFI	797
547	MVKQRCEKSSFYIFSSPSSQCNCYCLEVDL	576
798	EKLINVSRENGFKFNMMK..LQTSFPLSPSKFAKYGMDSVVEQNIVQDYCD	846
577	PGDEL RPSMQKLLQEKGLGGG..TDFPYECIDEWTKNKTHTVD	617
847	WIGISIDMKTLALMPNINLRIEGILCTNLNMOTKKASMWLKKKLSFLM	896
618	NIVILSDMHIAEGYS DINVRGSSIVNSI...KKYKDEVN	653
897	NNITHYFRKTTITDFANKTLNKLFISSGGYKYMCAKEYKD..HFKNLAM	945
654	PNIKIF...AVDLEGYK...KCLNLGDEFNENNYIKIFGM	687
946	SSHIDLEVSKIIYSVTRAFFKYLVCNIKDTIFGEEHYPDFFLSTLKHFI	995
688	SDSI...LKFISAKQGA...NMVE	706
996	IFSTKKYIFNRVC	1008
707	VI...KNFALQKIG	717

[illegible]

FIGURE 16

1 MEMDIDLODDIENL.....LPNTFNKYSSSCSDKKGCKTLKSGSKSPS... 42
491 TELAIKIAVNKNLDEIKGHTAIFSDVSGSHSTSHSGGAKKYGSVRTCLEC 540
43 LTIPKLQKO.....LEFYFSDANLYNDSFLRKLVLKSGEQRVEIETLL 85
541 ALVLGLMVKQRCESFYIFSSPSSQCNCYL.EVDLPDELRP SHQKLL 589

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 17

	Motif A	Motif B
Consensus	h--h h h---h--h	h---+-- h h---Sp
telomerase p123	GQPKLFFATMDIEKCYDSVNREKLSTFLKTTKLL-100	KFYRQTKGIPQGLCVSSILSSFYATLEESSLGFL
Dong (LINE)	KNRNLHCTYIDYKKAFDSPHSWLIQVLEIYKIN-	28-RQLAIKKGIYQGLSLSPJWFCALNPLSHQLWNR
al S.c.(group II)	FGGSNWFREVDLKKCFDTISHDLIIKELKRYISD-	26-HVPVGPVRCVQGLAPTSALCNAVLRLRRLAGLA
HIV-RT	LKKCKSVTVLDVGDAYFSVPLDEDFRKYTAFTIP-	7-GIRYQYNVLPQGWKGSIPAIFQSSMTKILXPFRKQN
L8543.12 y ² 11	VLPELYFMKFDVKSCYDSIPRMECMRILKDALKN-	68-KCYIREDEGLFQESSLSA?IVDLVYDDLLEFYSEPK

	Motif C	Motif D	Motif B
Consensus	h--YR DD hhh	Ch -h--- R	h-h Ch -h
telomerase p123	-14-LMR LT DDYLLITQENN-0-AVLFI E KLINVSR EN CFK FN TR LT -23-QDYCD WT GIS I		
Dong (LINE)	-16-HLIY MD IIKLYAKNDKE-0-MKKLID TT IFSN DIS MQ FL DR CK T-25-KC Y KY LV GF Q Q		
al S.c.(groupII)	-55-YVRY AD ILIGVLGSK N -2-KI IK RD LN FN LS .IG L TINE ER TL I -4-ET ? AR FL GY NT		
HIV-RT	-4-IYQY MD LYVGS HL EIG-1-HRT K IE EL RQ HL LR WL TP DR K HQ K-0-EP ? FL WN GY EL		
L8543.12	-8-IL KL AD DF LIISTD QQ Q.....VIN IK KL AM CFQ K YN AR AN R -41-IR SK SS SK G I ER		

1. *What is the purpose of this study?*

telomerase p43 LQKQLEEFYFSDANLYNDSFLRKLVLKSGEQRVEIETLLM
human La ICHQOEYFYFGDENLPRDKFLKEQI.KLDEGWVPLEIMIK
Xenopus LaA ICEQLEFYFGDHNLPDRDKFLKQOI.LLDDGWVPLETMIK
Drosophila La ILRQMEYFYGDANLNPRDKFLREIQIGKNEDGWVPLSVLVT
S. c. Lhp1p CLKQMEFYFSEFNFPYDFRLRTTAEK.NDGWVPISTIAT

[illegible]

THE
LIBRARY
OF THE
UNITED STATES
DEPARTMENT OF
COMMERCE
WASHINGTON
D. C.
20540

1 aactcattta attacttaalt taatcaacaa gattgataaa aagcagtaa taaaacccaa
61 tagatttaatt ttagaaagta tcaattgaaa aatggaaatt gaaaacaact aagcacaata
121 gccaaaagcc gaaaaattgt ggtgggaact tgaattagag atgcaagaaa accaaaatga
181 tatataagt aggggttaaga ttgacgatcc taagcaatat ctctgtaacg tcactgcagc
241 atgtttgttg taggaaggtta gttactacta agataaagat gaaaagaagat atatcatcac
301 taaagcactt cttaggttg ctgagtgta tcttgagttc atctgctagt tggcagtgta
361 catccgtaat gaactttaca tcagaactac cactaactac attgtagcat tttgtgtgt
421 ccacaagaat actcaacat tcatgaaaa gtacttcaac aaagcagtac tttgcctaa
481 tgacttactg gaagctgtg aatttgata ggttctctat attttgatg caactgaatt
541 caaaaatttg tatcttgata ggatacttct ataagataat cgtaaggaac tcactttccg
601 taagtgttta caaagatgag tcagaagcaa gttttctgaa ttcaacgaat actaacttgg
661 taagtattgc actgaatctt aacgtaagaa aacaatgttc cgttacctct cagttaccaa
721 caagtaaaag tgggattaaa ctaagaagaa gagaaaagag aatctcttaa ccaaacttta
781 ggcaataaag gaatctgaag ataagtccaa gagagaaact ggagacataa tgaacgttga
841 agatgcaatc aaggctttaa aaccagcagt tatgaagaaa atagccaaga gatagaatgc
901 catgaagaaa cacatgaagg cacctaaaat tcttaactct accttggaaat caaagtactt
961 gaccttcaag gatctcatta agttctgcca tatttctgag cctaaagaaa gagtctataa
1021 gatccttgggt aaaaaatacc ctaagaccga agagggaatc aaagcagcct ttgtgtattc
1081 tgcacttgca ccttcaatc ctgaattggc tggaaagcgt atgaagattg aaatctctaa
1141 aacatgggaa aatgaactca gtgcaaaagg caactctgt gaggtttggg ataatttaat
1201 ttcaagcaat taactcccat atatggccat gttacgtaac ttgtctaaca tcttaaaagc
1261 cgggttttca gatactacac actctattgt gatcaacaag atttgtgagc ccaaggccgt
1321 tgagaactcc aagatgttcc ctcttcaatt cttagtgcc attgaagctg ttaatgaagc
1381 agttactaag ggattcaagg ccaagaagag agaaaatatg aatcttaaag gtcaaatcga
1441 agcagtaaag gaagtgttg aaaaaaccga tgaagagaag aaagatatgg agttggagta
1501 aaccgaagaa ggagaatttg ttaaaagcaa cgaaggaatt ggcaagcaat acattaactc
1561 cattgaactt gcaatcaaga tagcagttaa caagaattta gatgaaatca aaggacacac
1621 tgcaatcttc tctgatgtt ctggttctat gactacatca atgtcaggtg gagccaagaa
1681 gtatgttcc gttctgactt gtctcgagtg tgcattagtc ctgtgtttga tggtaaaata
1741 acgttgtgaa aagtcctcat tctacatctt cagttacctt agtttcaat gcaataagt
1801 ttacttagaa gttgatctcc ctggagacga actccgtcct tctatgtaa aacttttga
1861 agagaagaa aaacttgggt gtgttactga ttccctcat gattgcattg atgaatggac
1921 aaagaataaa actcacgtag acaatatcgt tattttgtct gatgatgata ttgcagaagg
1981 atattcagat atcaatgtta gaggcagttc cattgttaac agcatcaaaa agtacaagga
2041 tgaagtaaat cctaacatta aaacttttgc agttgactta gaaggttacg gaaagtgcct
2101 taatctaggt gatgagttca atgaaaacaa ctacatcaag atattcggtg ttagcgattc
2161 aatcttaaag tcatctttag ccaagcaagg aggagcaaat atggtcgaag ttatcaaaaa
2221 ctgtgccctt caaaaaatag gacaaaagtg agtttttga gattcttcta taacaaaaat
2281 ctaccccccac tttttgtt tattgcatag ccattatgaa atttaaatla ttatctattt
2341 atttaagta ctacatagt ttatgtatcg cagtctatta gcctattcaa atgattctgc
2401 aaagaacaaa aaagattaaa a

FIGURE 20

MEIENNQAQQPKAEKLWWELELEMQENQNDIQVRVKIDDPKQYL
VNVTAACLLQEGSYQDKDERRYIITKALLEVAESDPEFICQLAVYIRNELYIRTTN
YIVAFCVVHKNTQPFIEKYFNKAVLLPNDLLEVCEFAQVLYIFDATEFKNLYLDRILS
QDIRKELTRKCLQRCVRSKFSEFNEYQLGKYCTESQRKKTMTFRYLSVTNKQKWDQTK
KKRKENLLTKLQAIKESEDKSKRETGDIMNVEDAIKALKPAVMKKIAKRQNAMKKHMK
APKIPNSTLESKYLTFKDLIKFCHISEPKERVYKJLGKKYPKTEEEYKAAFGDSASAP
FNPELAGKRMKIEISKTWENELSAKGNTAEVWDNLISSNQLPYMAMLRNLSNLIKAGV
SDTTHSIVTNKICEPKAVENSKMFPLQFFSAIEAVNEAVTKGFKAKKRENMNKKGQIE
AVKEVVEKTDEEKKDMELEQTEEGEFVKVNEGIGKQYINSIELAIKIAVNKNLDEIKG
HTAIFSDVSGSMSTMSGGAKKYGSVRTCLECALVLGLMVKQRCEKSSFYIFSSPSSQ
CNKCYLEVDPGDELPSMQKLLQEKGLGGGTDFPYECIDEWTKNKTHTVDNIVILSD
MMIAEGYSINVRGSSIVNSIKKYKDEVNPNIKIFAVDLEGYGKCLNLGDEFNENNYI
KIFGMSDSILKFISAKQGGANMVEVIKNFALQKIGQK

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 21

1 tcaatactat taattaataa ataaaaaaaa gcaaaactaca aagaaaatgt caaggcgtaa
 61 ctaaaaaaag ccataggctc ctataggcaa tgaacaaat ctgattttg tattacaaa
 121 tctagaagtt tacaaaagcc agattgagca ttataagacc tagtagtaat agatcaaaga
 181 ggaggatctc aagcttttaa agttcaaaaa ttaagattag gatggaaact ctggcaacga
 241 tgatgatgat gaagaaaaca actcaataa ataataagaa ttattaagga gagtcaatta
 301 gattaagtag caagttaaat tgataaaaaa agttgggtct aaggtagaga aagatttgaa
 361 ttgaacgaa gatgaaaaca aaaagaatgg actttctgaa tagcaagtga aagaagagta
 421 attaagaacg attactgaag aatagggtta gtattaaat ttagtattta acatggacta
 481 ccagttagat ttaaatgaga gtggtggcca tagaagacac agaagagaaa cagattatga
 541 tactgaaaaa tggtttgaaa tatctcatga ccaaaaaaat tatgtatcaa ttacgccaa
 601 ctaaaagaca tcatattgtt ggtggcttaa agattatttt aataaaaaaca attatgatca
 661 tcttaatgta agcattaaca gactagaaac tgaagccgaa ttctatgcct ttgatgatt
 721 ttcacaaaca atcaactta ctaataattc ttactagact gtaacatag acgttaatt
 781 tgataataat ctctgtatac tcgcattgct tagattttta ttactactag aaagattcaa
 841 tattttgaat ataagatctt ctataacaag aaattaatat aattttgaga aaattgggta
 901 gctacttgaa actatcttcg cagttgtctt ttctcatcgc cacttacaag gcattcatt
 961 acaagttcct tgcgaagcgt tctaatttt agttaactcc tcatcataaa ttacggttaa
 1021 agatagctaa ttataggtat actctttctc tacagactta aaattagttg acactaaca
 1081 agtccaagat tattttaagt tcttataaga attccctcgt ttgactcatg taagctagta
 1141 ggctatccca gttagtcta ctaacgctgt agagaacctc aatgttttac ttaaaaaggt
 1201 caagcatgct aatcttaatt tagtttctat ccctacctaa ttcaattttg atttctact
 1261 tgttaattta taacatttga aattagagtt tggattagaa ccaaatattt tgacaaaaca
 1321 aaagcttgaa aatctacttt tgagtataaa ataatacaaaa aatcttaaat tttaagatt
 1381 aaacttttac acctacgttg ctaagaacac ctccagaaaa cagatattaa aacaagctac
 1441 aacaatcaaa aatctcaaaa acaataaaaa tcaagaagaa actcctgaaa ctaaagatga
 1501 aactccaagc gaaagcacia gtggtatgaa atttttgat catctttctg aattaaccga
 1561 gcttgaagat ttacgcgita acttgaagc tacccaagaa atttatgata gcttgacaa
 1621 acttttgatt agatcaacaa atttaagaa gttcaaatga agttacaaat atgaatgga
 1681 aaagagtaaa atggatacat tcatagatct taagaatatt tatgaaacct taaacaatct
 1741 taaaagatgc tctgttaata tatcaaatcc tcatggaaac atttctatg aactgacaaa
 1801 taaagattct actttttata aatttaagct gaccttaaac taagaattat aacacgctaa
 1861 gtatactttt aagtagaacg aattttaatt taataacggt aaaagtgcac aaattgaatc
 1921 ttccāāta gaaagcttag aagatattga tagtctttgc aaatctattg ctcttgtaa
 1981 aaatttaca aatgttaata ttatcgccag ttgtctctat cccaacaata tttagaaaa
 2041 tccctcaat aagcccaatc ttctattttt caagcaattt gaataattga aaaatttgga
 2101 aaatgtatct atcaactgta ttctgatca gcatatactt aattctattt cagaattctt
 2161 agaaaagaat aaaaaataa aagcattcat ttgaaaaga tattatttat tacaatatta
 2221 tcttgattat actaaattat taaaacact tcaatagtta cctgaattaa attaagtta
 2281 cattaattag caattagaag aattgactgt gagtgaagta cataagtaag tatgggaaa
 2341 ccacaagcaa aaagctttct atgaaccatt atgtgagttt atcaagaat catcctaac
 2401 cctttagcta atagattttg accaaaacac tgaagtgtat gactctatta aaaagatttt
 2461 agaactata tctgagtcta agtatcatca ttatttgaga ttgaacccta gttatctag
 2521 cagtttaatt aaatctgaaa acgaagaat ttaagaactt ctcaaagctt gcgacgaaa
 2581 aggtgtttta gtaaaagcat actataaatt cctctatgt ttaccaactg gtacttatta
 2641 cgattacaat tcagatagat ggtgattaat taaatattag tttaataaa tattaaatat
 2701 tgaatattc ttgcttatt atttgaataa tacatacaat agtcattttt agtgttttga
 2761 atatatatta gttatttaatt tcatattttt aagtaataa ttattttica atcattttt
 2821 aaaaaatcg

FIGURE 22

MSRRNQKKPQAPIGNETNLDFVLQNLEVYKSQIEHYKTQQQIK
EEDLKLLKFKNQDQDGNNGNDDDEENNSNKQQELLRRVNQIKQQVQLIKKVGSKVEK
DLNLNEDENKKNGLSEQQVKEEQLRTITEEQVKYQNLVFNMDYQLDLNESGGHRRHRR
ETDYDTEKWFEISHDQKNYVSIYANQKTSYCWWLKDYFNKNNDHNLNVSINRLETEAE
FYAFDDFSQTIKLTNNSYQTVNIDVNFDDNLCILALLRFLSLERFNILNIRSSYTRN
QYNFEKIGELLETFVVFVSHRHLQGIHLQVPCEAFQYLVNSSSQISVKDSQLQVYSF
STDLKLVDTNKVQDYFKFLQEFPRLTHVSQQAIPVSATNAVENLNVLLKKVKHANLNL
VSIPTQFNFDYFVNLQHLKLEFGLPNILTKQKLENLLLSIKQSKNLKFLRLNFYTY
VAQETSRKQILKQATTIKNLKNNKNQEETPETKDETPSESTSGMKFFDHLSELTELED
FSVNLQATQEIYDSLHKLLIRSTNLKKFKLSYKYEMEKSKMDTFIDLKNIYETLNNLK
RCSVNISNPHGNISYELTNKDSTFYKFKLTLNQELQHAKYTFKQNEFQFNNVKSAKIE
SSSLESLEDIDSLCKSIASCKNLQNVNIIASLLYPNNIQKNPFNKPNLLFFKQFEQLK
NLENVSINCILDQHILNSISEFLEKNKKIKAFILKRYLLQYYLDYTKLFTLQQLPE
LNQVYINQQLLEELTVSEVHKQVWENHKQKAFYEPLCEFIKESSTLQLIDFDQNTVSD
DSIKKILESISESKYHHYLRNPSQSSSLIKSENEEIQELLKACDEKGVLVKAYYKFP
LCLPTGTYYDYNSDRW

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 23

MKILFEFIQDKLDIDLQTNSTYKENLKCGHFNGLDEILTTCFAL
PNSRKIALPCLPGDLISHKAVIDHCIIYLLTGELYNNVLTFGYKIARNEDVNNSLFCBS
ANVNVTLKGAAWKMFHSLVGTYAFVDLLINYTVIQFNGQFFTQIVGNRCNEPHLPPK
WVQRSSSSSATAAQIKQLTEPVTNKQFLHKLNNSSSFFPYSKILPSSSSIKKLTDLR
EAIPTNLVKIPQRLKVRNLTLQKLLKRHKRLNYVSILNSICPPLEGTVLDLSHLR
QSPKERVLFIIIVILQKLLPQEMFGSKKNKGKIIKNLNLLLSLPLNGYLPFDSLLKKL
RLKDFRWLFISDIWFTKHNFNENLNQLAICFISWLFRLPKIQTFFYCTEISSTVTI
VYFRHDTWNKLITPFIVEYFKTYLVENNVCRNHNSYTLNFNHNSKMRIIPKKSNEFR
IIAIPCRGADEEEFTIYKENHKNAIQPTQKILEYLRNKRPTSFTKIYSPTQIADRIKE
FKQRLLKKFNVLPELYFMKFDVKSCYDSIPRMECMRILKDALKNENGFFVRSQYFFN
TNTGVLKLFNVVNASRVKPYELYIDNVRTVHLSNQDVINVVEMEIFKTALWVEDKCY
IREDGLFQGSSLSAPIVDLVYDDLLEFYSEFKASPSQDTLILKLADDFLIISTDQQQV
INIKKLAMGGFQKYNANRDKJLA VSSQSDDDTVIQFCAMHIFVKELEVWKHSSTMN
NFHRSKSSKGIFRSLIALFNTRISYKTIDTNLSTNTVLMQIDHVVKNISECYKSAF
KDLSINVTQNMQFHSFLQRIIEMTVSGCPITKCDPLIEYEVRFITLNGFLESLSNTS
KFKDNIIILLRKEIQHLQAYIYIYIHIVN

LCVSYILSSFYYANLEENALQFLRKESMDPEKPETNLLMRLT
LCVSSILSSFYYATLESSLGLRDESMNPENPNVNLLMRLT

[illegible]

FIGURE 25

human
tez1
EST2
p123

Motif 0
AKFLHWLMSVYVVELLRSPFFYVTETTFQXNR
ISEIEWLVLGKRSNAXMCLSDFEKRRQIFAEFIYWLNSPIIPILQSFFYITESSDLNR
LKDFRWLFISD---IWFTKHNFFENLNQLAICFISWLPRLIPKIIQTFFYCTEISSVT-
TREISWMQVET-SAKHFFYYFDHEN-IYVLWKLLRWIFEDLVVSLIRCFYVTEQQKSYSK
* *

human
tez1
EST2
p123

Motif 1
LFFYRKSVHSLQSIGIRQHLKRVQLRDVSEAEVRQHREARPALLTSRLRPIPKP--DGL
TVYFRKDIWKLLCRPPI-TSMKMEAFEKINENNVRMDTQK-TTLPPAVIRLLPKK--NTP
IVYFRHDTWNKLITPFIYFKTYLVENNVCRNHNSYTLS--NPNHSMRIIPKKSNEF
TYYYRKNIWDVIMKMSI-ADLKKETLAEVQEKEVEWKKK-LGFAPGKLRLLPKK--TTF
* *

human
tez1
EST2
p123

Motif 2
RPIVNMHDYVVGARTFRREKRAERLTSRVKALP-SVLNYERA
RLITN-LRKRFLIKMGSNKKMLVSTNQTLRPVASILKHLINEESSGIPFNLEVYMKLLTF
RIIAIPCRGADEEEFTIYKENHKMAIQPTQKILEYLRNKRPTSFTKIYSPTQIADRIKEF
RPIMTFNKKIVNSDRKTTKLTNTKLLNSHMLKTLKN-RMFKDPPGPAVFNYDDVMKKY
* *

human
tez1
EST2
p123

Motif 3 (A)
KKDLLKXRMFOR-KKYFVRIDIKSCYDRIKQDLMFRIVKK-KLKDPEPVIRKYATIHATS
KQRLKKKFNNVLPelyfMKFDVKSCYDSIPRMECMRILKD-ALKNENGFFVRSQYFFNTN
EEFVCKWKQVGQPKLPFATMDIEKCYDSVNREKLSTFLKTTKLLSSDFWIMTAQILKRN
* *

FIGURE 26

ATTTATACTCATGAAAATCTTATTCGAGTTCATTCAAGACAAGCTTGACATTGATCTACA
 GACCAACAGTACTTACAAAGAAAATTTAAAATGTGGTCACTTCAATGGCCTCGATGAAAT
 TCTAACTACGTGTTTCGCACTACCAAATTCAGAAAAATAGCATTACCATGCCTTCCTGG
 TGACTIONAAGCCACAAAGCAGTCATTGATCACTGCATCATTTACCTGTTGACGGGCGAATT
 ATACAACAACGTACTAACATTTGGCTATAAAATAGCTAGAAATGAAGATGTCAACAATAG
 TCTTTTTTGGCATTCTGCAAATGTTAACGTTACGTTACTGAAAGGCGCTGCTTGAAAAAT
 GTTCCACAGTTTGGTCGGTACATACGCATTCGTTGATTTATTGATCAATTATACAGTAAT
 TCAATTTAATGGGCAGTTTTCCTCAAAATCGTGGGTAAACAGATGTAACGAACCTCATCT
 GCCGCCCAAATGGGTCCAACGATCATCCTCATCATCCGCAACTGCTGCGCAAATCAAACA
 ACTTACAGAACCAGTGACAAATAAACAATTTCTTACACAAGCTCAATATAAATTCCTCTTC
 TTTTTTTCCTTATAGCAAGATCCTTCCTTCATCATCATCTATCAAAAAGCTAACTGACTT
 GAGAGAAGCTATTTTTCCCAAAAATTTGGTTAAAATTCCTCAGAGACTAAAGGTACGAAT
 TAATTTGACGCTGCAAAAAGCTATTAAAGAGACATAAGCGTTTGAATTACGTTTCTATTTT
 GAATAGTATTTGCCCACCATTGGAAGGGACCGTATTGGACTTGTGCGATTGTAGTAGGCA
 ATCACCAAAGGAACGAGTCTTGAAATTTATCATTTGTTATTTTACAGAAGTTATTACCCCA
 AGAAATGTTTGGCTCAAAGAAAAATAAAGGAAAAATTATCAAGAATCTAAATCTTTTATT
 AAGTTTACCCTTAAATGGCTATTTACCATTTGATAGTTTGTGAAAAAGTTAAGATTAAA
 GGATTTTCGGTGGTTGTTTCAATTTCTGATATTTGGTTTACCAAGCACAAATTTGAAAACCTT
 GAATCAATTGGCGATTTGTTTCAATTTCTGGCTATTTAGACAATAATTCCTCAAAATTTAT
 ACAGACTTTTTTTTTACTGCACCGAAATATCTTCTACAGTGACAATTGTTTACTTTAGACA
 TGATACTTGAATAAACTTATCACCCCTTTTATCGTAGAATATTTTAAAGACGTACTTAGT
 CGAAAACAACGTATGTAGAAACCATAATAGTTACACGTTGTCCAATTTCAATCATAGCAA
 AATGAGGATTATACCAAAAAAAGTAATAATGAGTTCAGGATTATTGCCATCCCATGCAG
 AGGGGCAGACGAAGAAGAATTCACAATTTATAAGGAGAATCACAAAAATGCTATCCAGCC
 CACTCAAAAAATTTTGAATACCTAAGAAACAAAAGGCCGACTAGTTTTACTAAAATATA
 TTCTCCAACGCAAATAGCTGACCGTATCAAAGAATTTAAGCAGAGACTTTTAAAGAAAT
 TAATAATGTCTTACCAGAGCTTTATTTTCATGAAATTTGATGTCAAATCTTGCTATGATTC
 CATAACCAAGGATGGAATGTATGAGGATACTCAAGGATGCGCTAAAAAATGAAAATGGGT
 TTTTCGTTAGATCTCAATATTTCTTCAATACCAATACAGGTGTATTGAAGTTATTTAATGT
 TGTTAACGCTAGCAGAGTACCAAAACCTTATGAGCTATACATAGATAATGTGAGGACGGT
 TCATTTATCAAATCAGGATGTTATAAACGTTGTAGAGATGGAAATATTTTAAACAGCTTT
 GTGGGTTGAAGATAAGTGCTACATTAGAGAAGATGGTCTTTTTTCAGGGCTCTAGTTTATC
 TGCTCCGATCGTTGATTTGGTGTATGACGATCTTCTGGAGTTTTATAGCGAGTTTAAAGC
 CAGTCCTAGCCAGGACACATTAATTTTAAACTGGCTGACGATTTCTTATAATATCAAC
 AGACCAACAGCAAGTGATCAATATCAAAAAGCTTGCCATGGGCGGATTTCAAAAATATAA
 TGCGAAAGCCAATAGAGACAAAATTTAGCCGTAAGCTCCCAATCAGATGATGATACGGT
 TATTCAATTTTGTGCAATGCACATATTTGTTAAAGAATTGGAAGTTTGGAACATTCAAG
 CACAATGAATAATTTCCATATCCGTTGAAATCTAGTAAAGGGATATTTTGAAGTTTAAAT
 AGCGCTGTTTAACTAGAAATCTCTTATAAAAACAATTGACACAAAATTTAAATTCAACAAA
 CACCGTTCTCATGCAAATTGATCATGTTGTAAAGAACATTTCGGAATGTTATAAATCTGC
 TTTTAAAGGATCTATCAATTAATGTTACGCAAAATATGCAATTTTATTCTTCTTACAACG
 CATCATTGAAATGACAGTCAGCGGTTGTCCAATTACGAAATGTGATCCTTTAATCGAGTA
 TGAGGTACGATTACCATATTGAATGGATTTTTTGAAAGCCTATCTTCAAACACATCAA
 ATTTAAAGATAATATCATTCTTTTGAGAAAGGAAATTCACACTTGCAAGC

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

FIGURE 27

AKFLHWLMSVYVVELLRSFFYVTETTFQKNRLFFYRKSWSKLQSIGIRQHLKR
VQLRDVSEAEVRQHREARPALLTSRLRFIPKPDGLRPIVNMDYVVGARTFRREKR
AERLTSRVKALFSVLNYERA

bioRxiv preprint doi: <https://doi.org/10.1101/2021.03.10.434444>; this version posted March 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

FIGURE 28

[illegible]

[illegible]

FIGURE 30.

ggtagcgaattacttcttctcctaagcgaattgcttccctgaacgcctcctaattctctggaataattttacaagaactcaataacaataccaagtcgaattccaatatgaagg
 tgttattagtgatcgaataatttctatttctcggtcgttaccaagtataaggacaaaagaaacattccttccccctaaagacttttactttatattacitttcaaatatatttcg
 ggttcgcttacttttaacggtgactgttttagctgctacttctagccaaccgctgtttctaccccgctattggatagctctggagtagctcacagaaatccttacaaatctt
 ctgatgagactatattagattacacagtcggtcatattcttaacatggagccttacacttttagatgagtcacgctcgatgagtaggtattggatcatccaacgtttgccttg
 aaaaggttgataattattgcaaaatcatgcttagtggtggaatccgcgaaggtttttgatgcttgacacgcttagcatgattgagatattcaaaatttctatccactacaa
 ctcccttaacgcggttttttctatttcttctcatgttgttccaaatatgtatcatctctgtattaggctttttccggtttactctggaatcgtaccttttctacttccccctaatg
 aataatctaaattagtttcgcttataattgatagtagtagaaagattggtgattctactcgttaattgttattagtttaaagatactttgcaaacatttattagctatcattatataaaa
 aaaatcctataaattataaattataatcaatatttgcggtcactatttataaaacgttatgatcagtaggacactttgcatatatagttatgcttaattggttacttgaacttgcA T
 GACCGAACACCATACCCCCAAAAGCAGGATTCTTCGCTTTCTAGAGAATCAATATGTATACCTATGTAT
 CCTTAAATGATTATGTACAACCTTGTGAGAGGGTCGCCGGAAGCTCGTATAGCAATATATGCGAA
 CGCTTGAGAAGCGATGTACAAACGTCCTTTTCTATTTTTCTTCATTTCGACTGTAGTCGGCTTCGACAGT
 AAGCCAGATGAAGGTGTTCAATTTTTCTTCTCCAAAATGCTCACAGTCAGAGgtatataattttgtttgattttttctattcg
 ggatagctaatatatggcgagCTAATAGCGAATGTTGTAACAGATGTTTCGATGAAAGTTTTGAGCGTCGAAGGA
 ATCTACTGATGAAAGGGTTTTCCATGgttaaggatttcaattgtgaaatatttaccgtcaattactgttcaagagattgtattaaccgataaagAA
 TCATGAAGATTTTCGAGCCATGCATGTAAACGGAGTACAAAATGATCTCGTTTTCTACTTTTCTAATTA
 CCTTATATCTATACTTGAGTCAAAAAATTGGCAACTTTTGTAGAAAATgttaataccgggttaagattgtgcgcactttgaaca
 agactgacaagtatagTATCGGCAGTGTATGCCATGCATTACTTATTATCCAAAGGAAGTATTTTTGAGGCTCTTC
 CAAATGACAATTACCTTCAGATTTCTGGCATAACCACTTTTTAAAAATAATGTGTTTGAGGAACTGTGT
 CAAAAAAGAAAGCGAACCATTGAAACATCCATTACTCAAAATAAAAGCGCCCCGAAAGAAGTTTC
 CTGGAATAGCATTCAATTAGTAGGTTTAGCATTTTTTACAGGTCATCCTATAAGAAGTTTAAGCAAGgt
 aactaatactgttatccttcataactaatttagATCTATATTTTAACTTACACTCTATTTGTGATCGGAACACAGTACACATG
 TGGCTTCAATGGATTTTTCCAAGGCAATTTGGACTTATAAACGCATTTCAGTGAAGCAATTGCACAA
 AGTGATTCCACTGGTATCACAGAGTACAGTTGTGCCCAAACGTCCTAAAGGTATACCCTTTAATTGA
 ACAACAGCAAAGCGACTCCATCGTATTTCTCTATCAAAAGTTTACAACCATTATTGCCCATATATTGA
 CACCACGATGATGAAAAATCCTTAGTTATTCCTTAAAGCCGAACCAGGTGTTTGCCTTTCTTCGATC
 CATTCTTGTTTCGAGTGTTCCTAAATTAATCTGGGGTAACCAAAGGATATTTGAGATAATATTAAGGg
 tattgtataaaatttattaccactaacgattttaccagACCTCGAAACTTTCTTGAATTAATCGAGATACGAGTCTTTTAGTTTAC
 ATTATTTAATGAGTAACATAAAAGgtaatatgcaaaattttttaccattaattaacaatcagATTTTCAGAAATTGAATGGCTAGT
 CCTTGAAAAAGGTCAAAATGCGAAAAATGTGCTTAAGTGATTTTGAGAAACGCAAGCAAAATATTTGCGG
 AATTCATCTACTGGCTATACAATTCGTTTATAATACCTATTTTACAATCTTTTTTTTATATCACTGAATC
 AAGTGATTTACGAAATCGAACTGTTTATTTTAGAAAAGATATTTGGAAACTCTTGTGCCGACCCTTTAT
 TACATCAATGAAAATGGAAGCGTTTGAAAAATAAACGAGgtattttaagattttttgcaaaaagcctaattttcagAACAA
 TGTTAGGATGGATACTCAGAAAACACTTTGCCTCCAGCAGTTATTCGTCTATTACCTAAGAAGAATAC
 CTTTCGTCTCATTACGAATTTAAGAAAAAGATTCTTAATAAAAGgtatttaattttgtcatcaatgtacttacttctaattctattattag
 cagATGGGTTCAAAACAAAAAATGTTAGTCAGTACGAACCAAACTTTACGACCTGTGGCATCGATACTG
 AAACATTTAATCAATGAAGAAAGTAGTGGTATTCATTTAACTTGGAGGTTTACATGAAGCTTCTTACT
 TTTAAGAAGGATCTTCTTAAGCACCGAATGTTTGGgttaattataatgcgagattcctcattattaattngcagGCGTAAGAAG
 TATTTTGTACGGATAGATATAAAATCCTGTTATGATCGAATAAAGCAAGATTTGATGTTTCGGATTGTT
 AAAAAAGAACTCAAGGATCCCGAATTTGTAATTCGAAAGTATGCAACCATACATGCAACAAGTGACCG
 AGCTACAAAAAACTTTGTTAGTGAGGCGTTTTCTATTgttaagttatttttcttggaatttttaacaaattccttttagTTGATAT
 GGTGCCTTTTGAAGAGTCGTGCAGTTACTTTCTATGAAAACATCAGATACTTTGTTTGTGATTTTGT
 GGATTATTGGACCAAAAGTTCTTCTGAAATTTTTAAATGCTCAAGGAACATCTCTCGGACACATTGT
 TAAAGgtataccaattgtgaattgaataacactaatgaactagATAGGAAATTCTCAATACCTTCAAAAAGTTGGTATCCCTC
 AGGGCTCAATTCTGTCTATCTTTTTGTGTCTTTCTATATGGAAGATTTGATTGATGAATACCTATCGTT
 TACGAAAAAAGAAAGGATCAGTGTGTTACGAGTAGTCGACGATTTCTCTTTATAACAGTTAATAAAA
 AGGATGCAAAAAAATTTTGAATTTATCTTTAAGAGgtgagttgtctgctccttaagttctaaccgttgaagGATTTGAGAA
 ACACAATTTTTCTACGAGCCTGGAGAAAACAGTAATAAACTTTGAAAATAGTAATGGGATAATAAACA
 ATACTTTTTTAAATGAAAGCAAGAAAAGATGCCATTCTTCGGTTTCTCTGTGAACATGAGGTCTCTTG
 ATACATTGTTAGCATGTCCTAAAATTGATGAAGCCTTATTTAACTCTACATCTGTAGAGCTGACGAAAC
 ATATGGGGAAATCTTTTTTTTACAAAATTCTAAGgtatactgtgtaactgaataatagctgacaaataatcagATCGAGCCTTGC
 ATCCTTTGCACAAGTATTTATTGACATTACCCACAATTCAAAATTCAATTCTTGCTGCAATATATATAG
 GCTAGGATACTCTATGTGTATGAGAGCACAAGCATACTTAAAAAGGATGAAGGATATATTTATCCCC
 AAAGAATGTTTCATAACGGgtgagttacttatttaactagaaaagtcattaattaaccttagATCTTTTGAATGTTATTGGAAGAAAA
 ATTTGGAAGAAAGTTGGCCGAAATATTAGGATATACGAGTAGGCGTTTCTTGTCTCTGCAGAAGTCAA

1000 900 800 700 600 500 400 300 200 100 0

FIGURE 30 (cont.)

[illegible]

[illegible]

EST2 pep	FFYCTEISST	VTIVYFRHDT	WN----	KLIT	P-----	FIVE	YFK--TYLVEN	40
Euplotes pep	FFYVTEQOKS	YSKTYYYRKN	IWDVI--MKMS	IAD----	LKK	ETLA--EVQE		43
Trans of tetrahymen	-----KHKE	GSQIFYRKP	IWKLVSKLTI	VKVRIQFSEK	NKQMKNNFYQ			44
Consensus	FFY.TE..K.	.S..YYRK.	IW...-KL..F..KV..			50
EST2 pep	NVCRNHNSY-	-----	TLSNFNHSM	FLTPKKSNE	FRITIAIPCRG			79
Euplotes pep	KEVEEWKKS	L-----	GFAPCKL	FLTPKCKIT	FRFIMTFNKK			78
Trans of tetrahymen	KIQLEENLE	KVEEKLIPE	SFOKYPOCK	FLTPKCKGS	FRFIMTFLRK			92
Consensus	K...E....	-----F..GK	FLTPKCK	FRFIMTF..RK			100
EST2 pep	ADEEBFTIYK	ENHNAIQPT	QKILEYRKN	RPTSFTKIYS	PTQIADRIKE			129
Euplotes pep	IVNSDRKTTK	LTTNTKLNS	HLMLKTKN	-----RMFK	DPFGFAVFN			120
Trans of tetrahymen	DKQKNIK---	LNLNQILMS	OLVFRNKK	-----ML-G	QKIGYSVFD			130
ConsensusK..K	LN.N..L..S	QL.L..LKN	-----	...IG..VF			150
EST2 pep	FKQRLLKFN	NVL-----	PELYFMKFD	VKSCYD				157
Euplotes pep	YD-DVMKKYE	EFVCKWKQVG	QPKLFFATMD	IEKCYD				155
Trans of tetrahymen	NK-QISEKFA	QFIEKWKNGK	PELYYVTL-	----				158
Consensus	.K-...KKF.	.F..KWK..G	ELLYF.T.D	...CYD				186

FIGURE 32

S-1: FFY VTE TTF QKN RLF FYR KSV WSK
S-2: RQH LKR VQL RDV SEA EVR QHR EA
S-3: AFT FRR EKR AER LTS RVK ALF SVL NYE

A-1: AKF LHW LMS VYV VEL LRS FFY VTE TTF Q
A-2: LFF YRK SVW SKL QSI GIR QHL KRV QLR DVS
A-3: PAL LTS RLR FIP KPD GLR PIV NMD YVV

103000 3232000

FIGURE 33

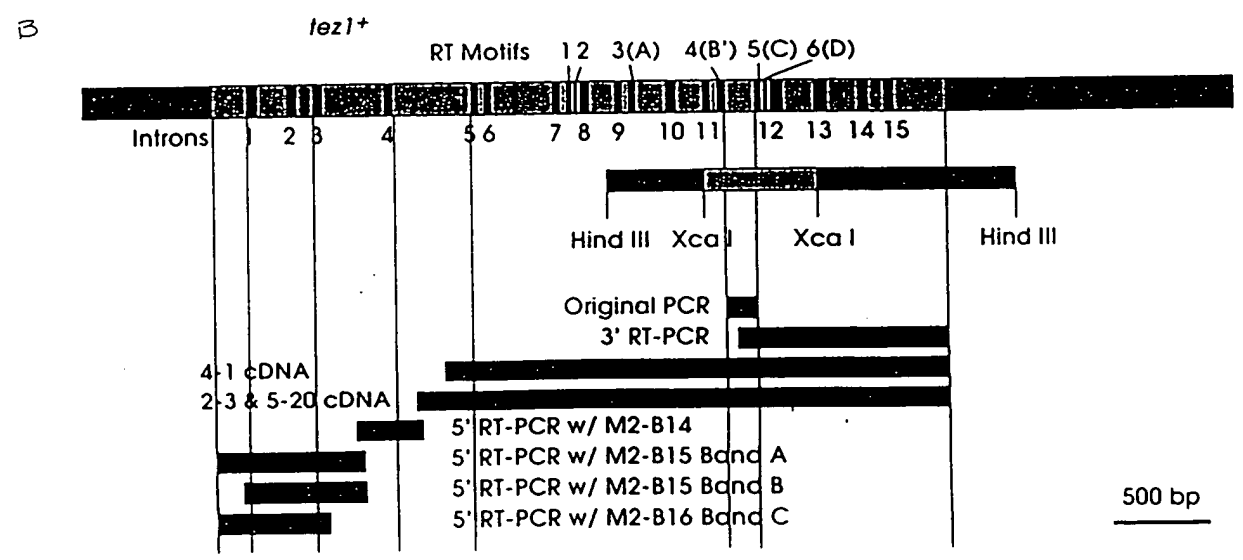
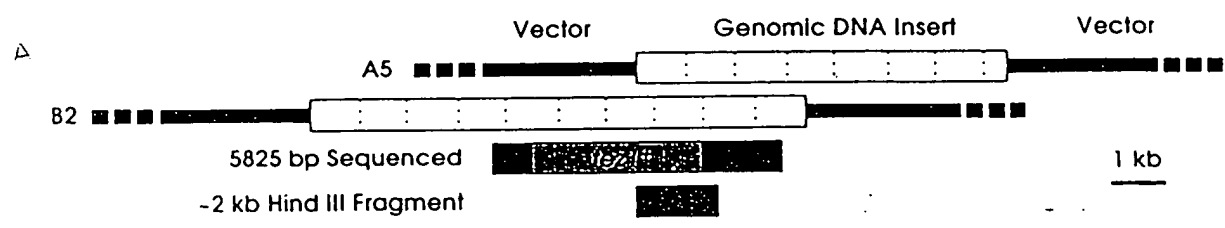


FIGURE 34

Poly 4

		t		t		c		
	t	a	a	g	c	c	t	c g
5'-	cag	acc	aaa	gga	att	cca	taa	gg -3'
	Q	T	K	G	I	P	Q	G

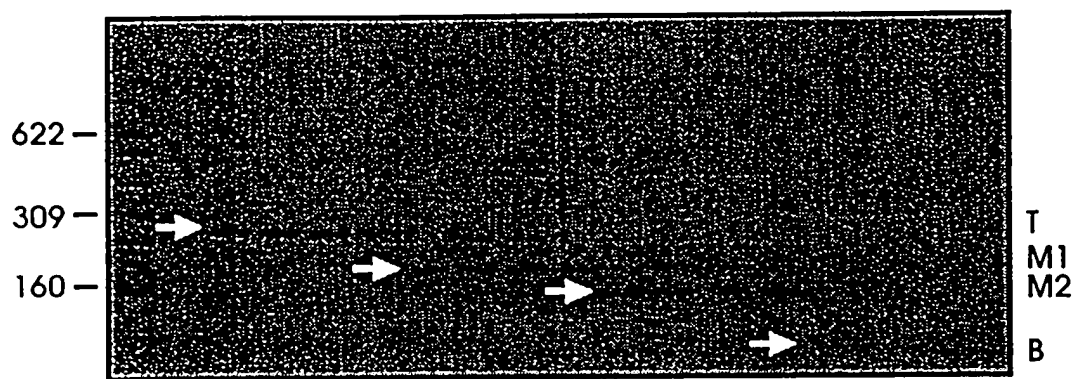
4 (B')

5 (c')

	D	D	Y	L	L	I	T	
3'-	ctg	ctg	atg	gag	gag	tag	tgg	-5'
	a	a	a	a	a	a	a	
				t	t	t	t	
						c	c	
						<u>Poly 1</u>		

105500 32521500

FIGURE 35



Motif B' (4)
QTKGIPQG

Motif C (5)
DDYLLIT

FIGURE 36

PCR Product M2 showed Reasonable Match with Other Telomerase Proteins

Ot	LCVSYILSSFYYANLEENALQFLRKESMDPEKPETNLLMRLT
Ea_p123	KGIPQGLCVSSILSSFYYATLEESSLGFLRDESMNPENPNVNLLMRLTDDYLLIT
Sp_M2	SILSSFLCHFYMEDLIDEYLSFTKKK-----GSVLLRVV
Sc_p103	DGLFQGSSLSAPIVDLVYDDLLEFYSEFKASPS-----QDTLILKLADDFLIIS
	. * . * . *

Q K V G I P Q G
gaa aaa gtt ggt atc cct cag gg..... <---Actual Genomic Sequence.

Poly 4

t t c
a a g c c t c g
Cag acc aaa gga att cca taa gg ---->
ag acc aaa gga att cca tca ggC TCA ATT CTG TCA TCT TTT TTG TGT CAT TTC TAT ATG
tc tgg ttt cct taa ggt agt ccG AGT TAA GAC AGT AGA AAA AAC ACA GTA AAG ATA TAC
K G I P S G S I L S S F L C H F Y M

GAA GAT TTG ATT GAT GAA TAC CTA TCG TTT ACG AAA AAG AAA GGA TCA GTG TTG TTA CGA
CTT CTA AAC TAA CTA CTT ATG GAT AGC AAA TGC TTT TTC TTT CCT AGT CAC AAC AAT GCT
E D L I D E Y L S F T K K K G S V L L R

GTA GTC gac gac tac ctc ctc atc acc
CAT CAG ctg ctg atg gag gag tag tgg.

V V D D Y L L I T

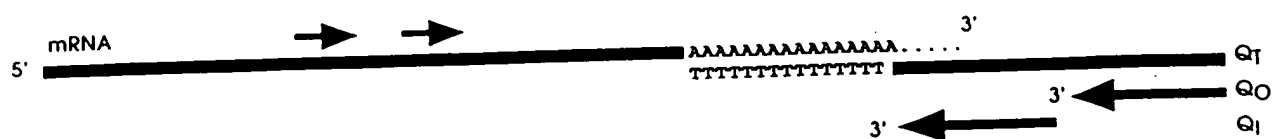
<---- ctg ctg atg gag gag tag tgg
a a a a a a a a
t t t t
c c

Poly 1

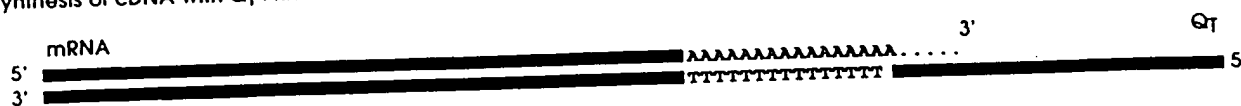
.....gac gat ttc ctc ttt ata aca..... <---Actual Genomic Sequence.
D D F L F I T

FIGURE 37

3' RT PCR Strategy



1. Synthesis of cDNA with Q_T Primer.



2. First Round PCR Using Outside Primer and Q_O Primer.



3. Second Round PCR Using Inside Primer and Q_I Primer.



4. Sequence Second Round PCR Products Using Inside Primer or Q_I Primer.

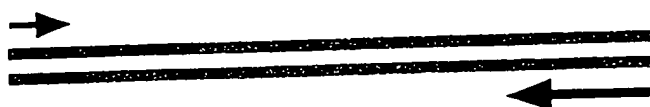


FIGURE 38

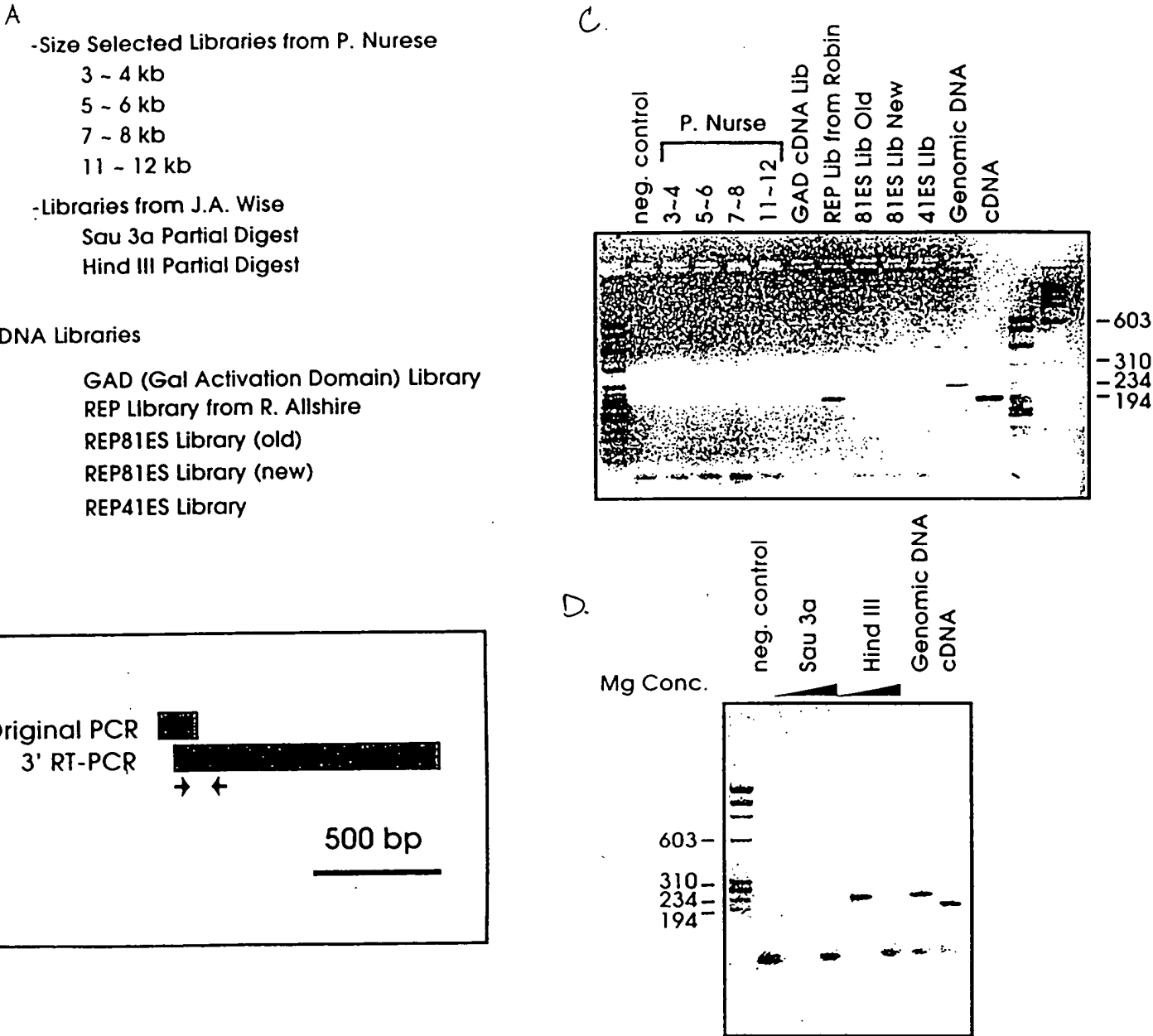


FIGURE 39

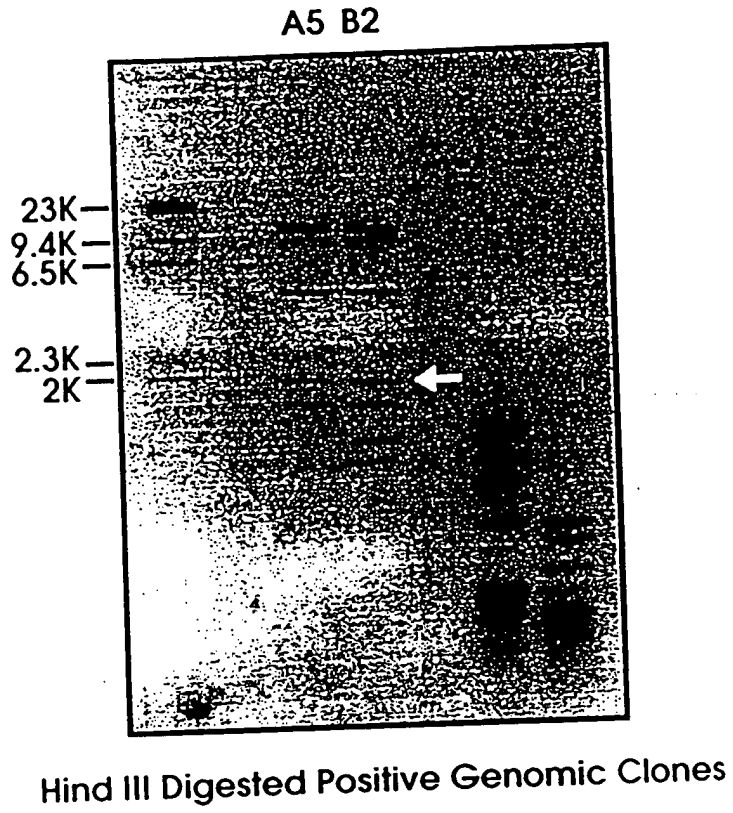
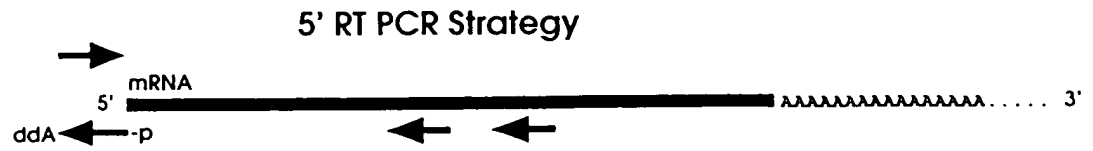
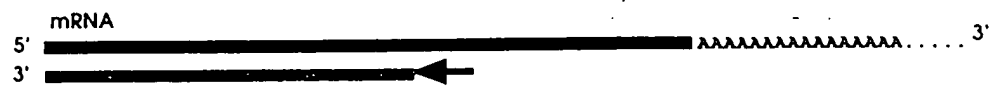


FIGURE 40



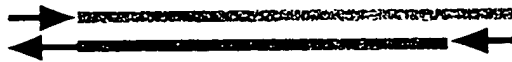
1. Synthesis of cDNA with Specific Downstream Primer.



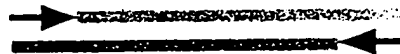
2. Ligate Oligo with 5'-P and blocked 3' to cDNA using T4 RNA Ligase.



3. First Round PCR



4. Second Round PCR

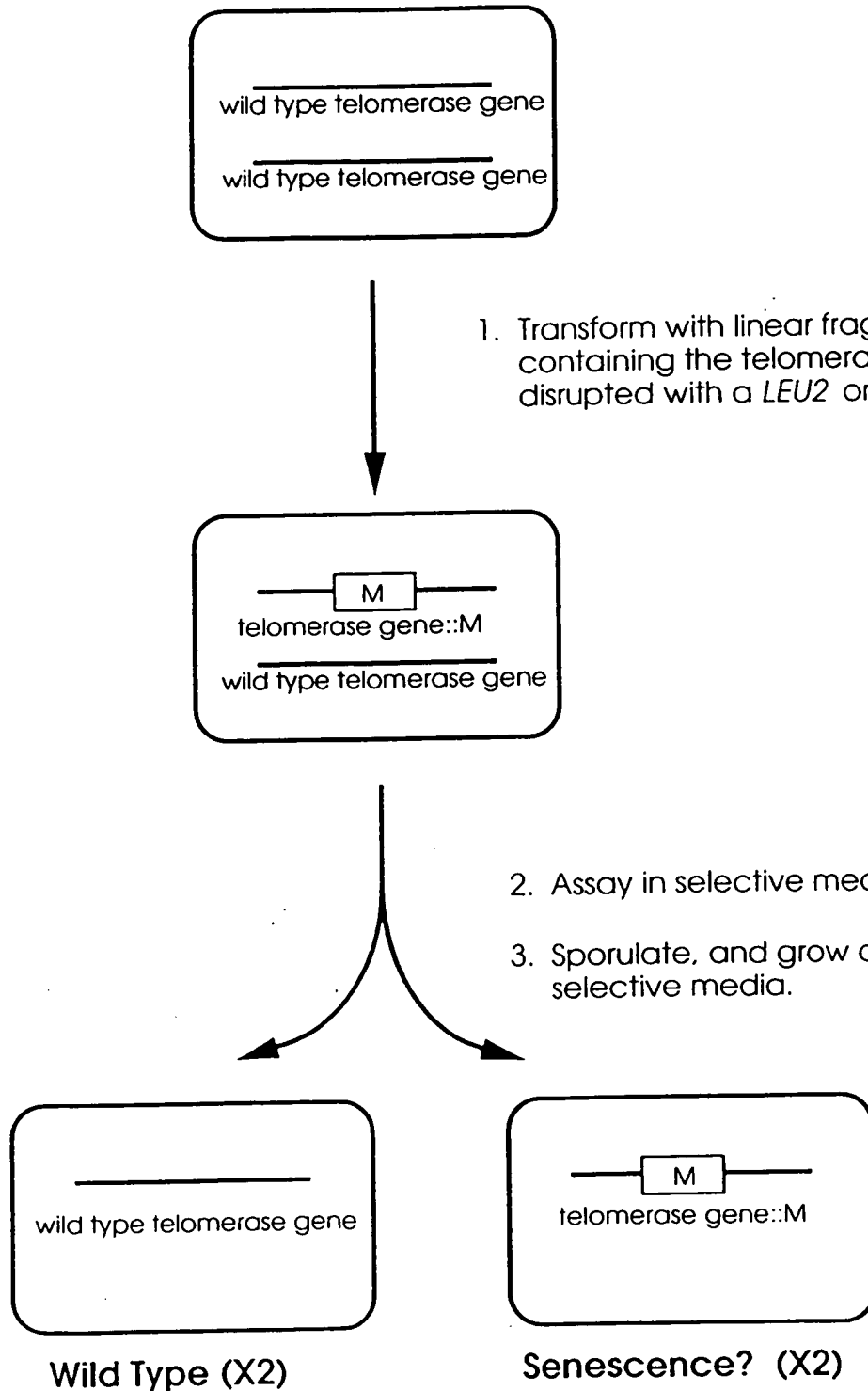


[illegible][illegible]

Sp_Trip 1MTEHMTKPSRIILRFLENOVYVLTCT 2
Sc_East2 1MKLFFET 2
Ea_p122 1 MEVDYDNOAGAHQHGMSALKTYCEEIKKATLVTSF 2
Sp_Trip 24 LDYDYLVLRSQPA.....ZNICERLSEVOTFSF 27
Sc_East2 24 IODKLDIDLOLTH.....ENLKKCGHFNGLD 24
Ea_p123 24 IOKVIRCANQSO.....H.....DLEDICIFAOTN 24
Sp_Trip 34 IFILNATYVVOFSKPDQGVQVSPSCSOSEL 34
Sc_East2 34ITTCFALPHSR-KIALPCLOGLSHKAY 34
Ea_p123 34 IVATPROYNEEDFKVIARKEVFSFTOLMTE 34
Sp_Trip 51 CVIKMFDESFEERR-LLNKK.....FNMNEDFRAMH 57
Sc_East2 51 CYYLLDTDELTV.....NHTYVTKIARNED 57
Ea_p123 51 CLVGLLSSSOVSROKLOCF.....FOLKNGD 57
Sp_Trip 123 VGVONDVLVFTFFPHYLISILESKN.....OLLEII 126
Sc_East2 123VHNSLFCFSAHNSVYTLKGLSFMFSL 126
Ea_p123 123LAKTHLTLTALSTKQGVFFOOD.....NOVRMI 126
Sp_Trip 126 SDAMMY.....LSKOBIFEALPNHNL.....ISLPLFKN 126
Sc_East2 126 TYAFYD.....LINTTYVIO-GOFFT.....VWRCEHP 126
Ea_p123 126 NELFRN.....TYKYLIFORTSEOTLV.....PCNNMFVOM 126
Sp_Trip 129 NVFEETVSKKREKTIETSIOTN.....ESAKKE 216
Sc_East2 129 LLPPKWVO.....RBSSSATAAOI.....KOLTER 216
Ea_p123 129 KLVKDFDK-KOKGQAADHOPRECCSTCYKN 216
Sp_Trip 216 WMSISISRFISIFVRSSTYKFKODLVFNLSHSD 216
Sc_East2 216KOFKLNHNSISSFFP 216
Ea_p123 216 NEK.....DHFLNNHNVPMHNNMSRTVIFCTYFHN 216
Sp_Trip 272 RNTVHMWLOVIFPROFOLINAFQVOLKHEVILP 272
Sc_East2 272YKKILPSS.....SIKKLTOLREAF 272
Ea_p123 272NOHFFKKHEFYNNHNISAMORAOTI 272
Sp_Trip 280 VS.....OSTVVPKRLKXYPLIOTAKRLHLS 213
Sc_East2 280 TH.....LILIPORNIYHNLKLOLHMRHLR 213
Ea_p123 280 PTNIFRFRHRIKKLKOKVIEKIAYMLEKVKOFN 280
Sp_Trip 314 LSKVYHNY.....YID-THODDERISVSLKPNQ 314
Sc_East2 314 VYSILNS.....PLEQTVLHNSVYTLKGLSFMFSL 314
Ea_p123 314 PHVYLTKE.....LPMWREKOKIENLHNTREEK 314
Sp_Trip 343FALRSILVRVYFKLI 343
Sc_East2 343LKITVILOKLLQEM 343
Ea_p123 343 SKYVEELPSYTTOMKC.....VHEFFYHILKOD 343
Sp_Trip 360 WGNORFEIIL.....DLETFLK.....RYSESFLNLS 360
Sc_East2 360 FOSKKNKEIIL.....LLNLLLS.....PLNGLVFPDS 360
Ea_p123 360 LKIG-RRNRKFO.....KRVKYYE.....KNHELKNHLE 360
Sp_Trip 373 NIKSIEEILVLQKKNAMCMLSDP.....KKKOFIA 429
Sc_East2 373 KLRKLOKFLFIS.....DIWFYKHF.....MLNOLAI 429
Ea_p123 373 KINTREISNOVETS-ARKEFFYHFN.....N-IVYLV 429
Sp_Trip 430 ESIY.....LYNSFVILPILOS.....SSDLNRHTY 430
Sc_East2 430 CFISILFROLIPIKIOT.....CISSTYTV 430
Ea_p123 430 KLLRIFEDLVVSLRC.....V.....OOKSYSKTY 430
Sp_Trip 460 FKDOI.....KLLCRVPLTSMKNSEAFKINENHVRMD 460
Sc_East2 460 FKDOI.....KLLCRVPLTSMKNSEAFKINENHVRMD 460
Ea_p123 460 YKKHIDVIMKNSADLKKETLAEVQEKREVEW 460
Sp_Trip 466 TOKTTLPPAVI.....LL.....NT.....LTHMRKFL 466
Sc_East2 466 YTLNFWHNSK.....L.....TT.....PMFTFKKKIV 466
Ea_p123 466 EKSGLOFAPOK.....L.....TT.....PMFTFKKKIV 466
Sp_Trip 570 KESKKNKMLVSTHOLRPVASI.....KHIE..... 562
Sc_East2 570 EMOE-PTTIEENKHAIOPTOKI.....YKRNMF 562
Ea_p123 570 MSD-REKTKTLTYTNKLLNSMLK.....YKRNMF 562
Sp_Trip 583 ESSOIFPNLEVMYKMLTKFKDL.....RMFOR-KK 566
Sc_East2 583 SFTKITSPTIOADRIRKEFORALL.....KPNHVPPEL 566
Ea_p123 583 KOPPOFAPVHYTODMYKKEFFYC.....KNOGQPL 566
Sp_Trip 586 YVRYIKS.....RIKODLFRVIL.....KKLKDPE 616
Sc_East2 586 YMKFVKYS.....SIPMECHKRI.....DALKNENG 616
Ea_p123 586 FEATRIER.....SVNHEKLSYFL.....TTKLLSSD 616
Sp_Trip 617 FRYKVIATINATEDRAN..... 626
Sc_East2 617 FVRSQVFFHTNY..... 626
Ea_p123 617 WYMTAIDLRKKNHIVIDSKNFRKEMEDYFROK 626
Sp_Trip 626 FYSEAFSYVDMFVFF.....VOLLS-.....MKTSD 616
Sc_East2 626VLFKFN.....VHNSR-.....PPKPYE 616
Ea_p123 626 FOXIALEGGOVPTFL.....LEWENODLNAKKT 616
Sp_Trip 686 DFYDVTYKSSSEIFPKMLKHLNSGVHIGNSO 676
Sc_East2 686 DNYTNYSHOQVINYVNEHIFKTLAWVEDK 676
Ea_p123 686 EKQKRYNKKQMDLLOLPVINYCQNTVYFNFOK 676
Sp_Trip 689 LQKVIP.....SIL.....SPLCHFYMED.....IDEVLSFK 721
Sc_East2 689 IREDFK.....SSLAPV.....VYVYVIEDEKKA 721
Ea_p123 689 KOTKIPILCLV.....SILSSFYVATRESSLGLR 721
Sp_Trip 726 KKG.....SYLLRVN.....F.....FVYNNKODAKK 726
Sc_East2 726 SPSOD.....TLILKLLA.....F.....FVYNNKODAKK 726
Ea_p123 726 DESMHPENPHVHLMLT.....YLLTTOENHVL 726
Sp_Trip 762 FLNLRLQFKEMFSLT.....TVYNFN.....HG..... 716
Sc_East2 762 ICKLAMDQFOKYHAKANRO.....ILAVSOD..... 716
Ea_p123 762 FIKLILNYSRSEKQFPHNM.....LOTSFPL.....PSKFA 716
Sp_Trip 767IINISSEKKKMPFPFOSVNMNSLDLTL 716
Sc_East2 767DDTVIOFCA.....MHIFVKELEVHNSMT 716
Ea_p123 767 KYAMDVSYEONIVODDYVYKIGMSDKETLAL 716
Sp_Trip 811 ACPKIDAEALFNSTSVLTKHMSGFYVILRSS 716
Sc_East2 811 NNFHIRESSKQIFRSLALFNTRISYITDTH 716
Ea_p123 811 NMLRIEIGLCTLNLNMTKCAKSMWLKLEFS 716
Sp_Trip 886 ASFAGVYFIDTHNNSFSECCNIVYLGYSMCMR 716
Sc_East2 886 DNYTNYSHOQVINYVNEHIFKTLAWVEDK 716
Ea_p123 886 MNNITHYFRTKLTTEOFANLTKVLFISQGT 716
Sp_Trip 906 AGAYLKR.....IFKORNFITOLLNVKLEWKK 716
Sc_East2 906YKSA.....VYVYVIEDEKKA 716
Ea_p123 906 YMOCAKEY.....NFKKILAMSMIDLEVSKIYEM 716
Sp_Trip 916 LAEILQYTSRRFLSSAEVHFLCLOMRDCE 916
Sc_East2 916 TVSQDPIYKCDPLKEDYFGEENTYPOFFLSTKHF 916
Ea_p123 916 TRAPPEYLVKILQYTSRRFLSSAEVHFLCLOMRDCE 916
Sp_Trip 966 FTYNPFCEOLIVOT.....OSTLDILKPLRV.....ROVLF 966
Sc_East2 966 TS.....KONHILKREIGH.....QATY 966
Ea_p123 966 IEIFS.....TKKY.....NRYCMILKAKEAK.....KSDOC 966
Sp_Trip 986 LHRRIAD- 986
Sc_East2 986 IYINIVH 986
Ea_p123 986 OSILOYDA 986

FIGURE 43

Disruption strategy for the putative telomerase genes.



(These cells will show a senescence phenotype if the disrupted gene encodes a telomerase subunit.)

FIGURE 44

An Example of Confirmation of *tez1* disruption By PCR

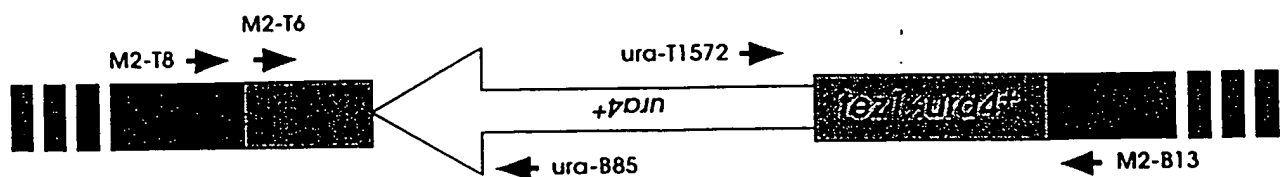
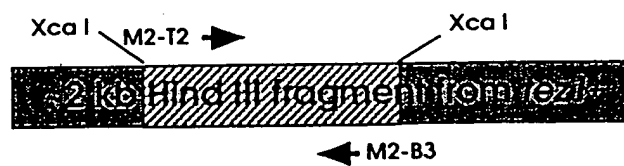
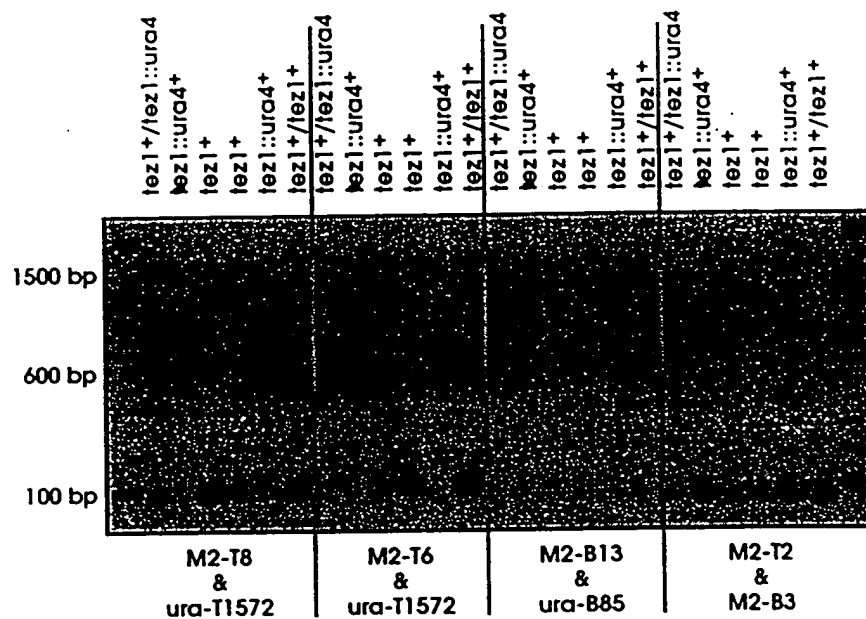


FIGURE 45

Tez1 disruption causes progressive shortening of telomeres in *S. pombe*

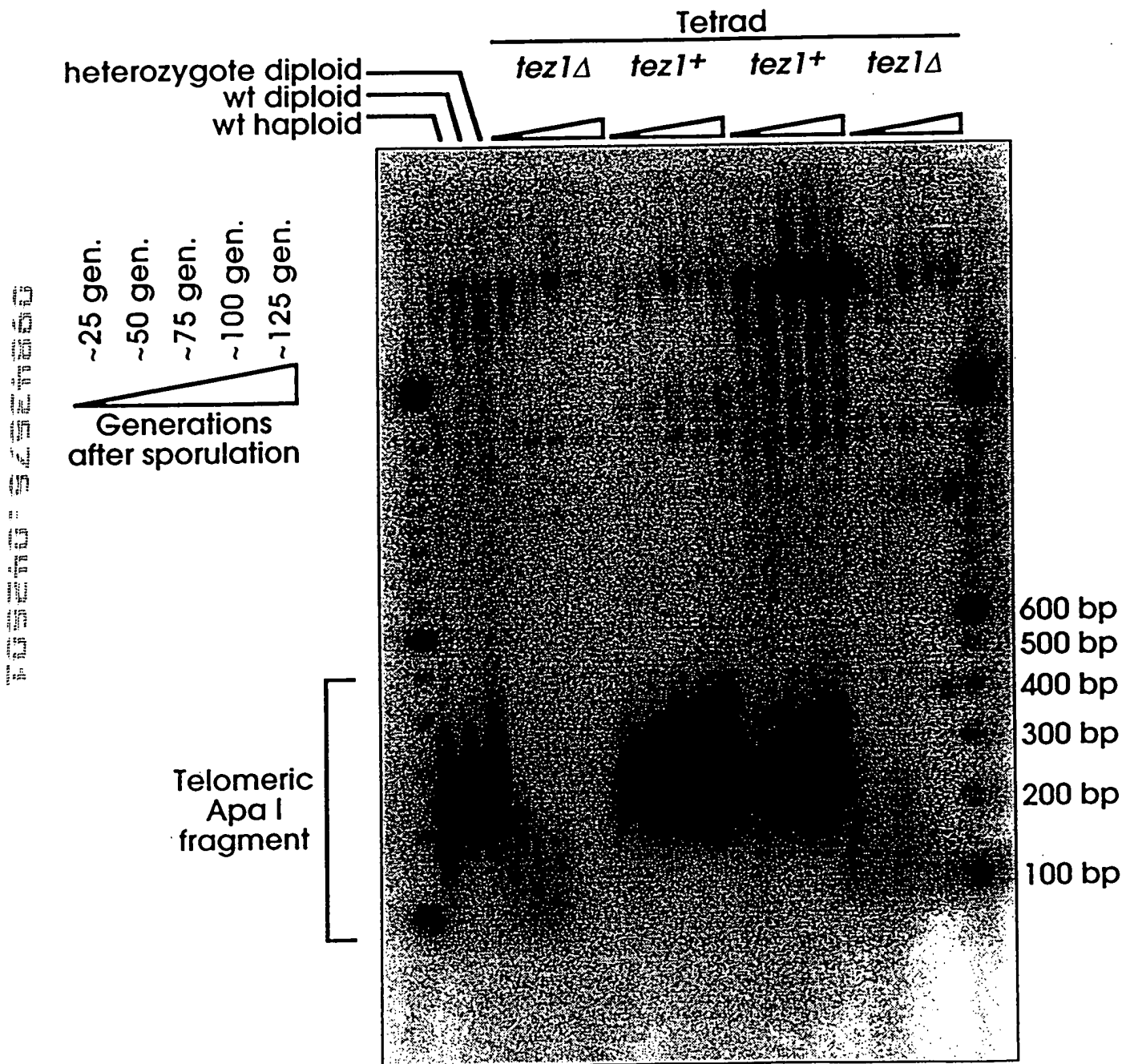


FIGURE 46

1 ggtaccgatttactttcctttcttcataagctaattgcttcctcgaacgctcctaaatctctggaaatatttttacaaga 80
 81 actcaataacaataccaagtcaaattccaatatgaagggtgttattagtgatcgataaatatttctattttatcggtcgta 160
 161 ccaagtataaggacaaaaagaacaacttccttccccctaaagacttttactttattaatttacttttcaaataatatttcg 240
 241 ggttcgcttacttttaacgtggtactgttttagctgctacttctagccaaccgctgtttctaccccgctcattggatat 320
 321 agctcttggagtagctcacagaaatccttacaaatcttctgatgagactatattagattcattacagtcggtgcatattc 400
 401 ttaacatggagccttacacttttagatgagtcacgtcgcatgatggagtatttgggtatcatccaacgtttgccttgaaaag 480
 481 gttgataattatttgcaaaatcatgtccttagtggtggttaatccgcgaaagttttttgatgcttgcacacgtctagcatg 560
 561 attgagatattcaaaaatttctatccactacaactcctttaacgcggttttatttttctattttctatttctcatgttgtt 640
 641 ccaaataatgtatcatctcgtattagggttttttccgttttactcctggaatcgtaccttttctactattccccctaata 720
 721 ataataaaattagtttgcgttataattgatagtagtagaaagattgggtgattctactcgtgtaattgtattagtttaaa 800
 801 gatactttgcaaaacatttattagctatcattatataaaaaaaatcctataattataaatattaatcaatatttgcggtc 880
 881 actatttattttaaacggttatgatcagtaggacactttgcatatatatagttatgcttaattggttacttgtaacttgc 958

 959 ATG ACC GAA CAC CAT ACC CCC AAA AGC AGG ATT CTT CGC TTT CTA GAG AAT CAA TAT GTA 1018
 1 M T E H H T P K S R I L R F L E N Q Y V 20

 1019 TAC CTA TGT ACC TTA AAT GAT TAT GTA CAA CTT GTT TTG AGA GGG TCG CCG GCA AGC TCG 1078
 21 Y L C T L N D Y V Q L V L R G S P A S S 40
 1079 TAT AGC AAT ATA TGC GAA CGC TTG AGA AGC GAT GTA CAA ACG TCC TTT TCT ATT TTT CTT 1138
 41 Y S N I C E R L R S D V Q T S F S I F L 60
 1139 CAT TCG ACT GTA GTC GGC TTC GAC AGT AAG CCA GAT GAA GGT GTT CAA TTT TCT TCT CCA 1198
 61 H S T V V G F D S K P D E G V Q F S S P 80

FIGURE 46 (cont.)

1199	AAA	TGC	TCA	CAG	TCA	GAG	gtatatatatat	ttttt	gtttt	gatttttt	ctattc	gggatag	ctaata	tatat	gggcag	1272					
81	K	C	S	Q	S	E										86					
1273	CTA	ATA	GCG	AAT	GTT	GTA	AAA	CAG	ATG	TTC	GAT	GAA	AGT	TTT	GAG	CGT	CGA	AGG	AAT	CTA	1332
87	L	I	A	N	V	V	K	Q	M	F	D	E	S	F	E	R	R	R	N	L	106
1333	CTG	ATG	AAA	GGG	TTT	TCC	ATG	gtaagg	tattct	aattgt	gaaatatt	tacctg	caattact	gtttc	aaagaga	1405					
107	L	M	K	G	F	S	M									113					
1406	ttgtatttta	aaccgataaa	g	AAT	CAT	GAA	GAT	TTT	CGA	GCC	ATG	CAT	GTA	AAC	GGA	GTA	CAA	AAT	1469		
114				N	H	E	D	F	R	A	M	H	V	N	G	V	Q	N	128		
1470	GAT	CTC	GTT	TCT	ACT	TTT	CCT	AAT	TAC	CTT	ATA	TCT	ATA	CTT	GAG	TCA	AAA	AAT	TGG	CAA	1529
129	D	L	V	S	T	F	P	N	Y	L	I	S	I	L	E	S	K	N	W	Q	148
1530	CTT	TTG	TTA	GAA	AT	gtaaata	accggtta	agatgtt	gcgcacttt	gaacaag	actgaca	agtatag	T	ATC	GGC	1601					
149	L	L	L	E	I								I	G		155					
1602	AGT	GAT	GCC	ATG	CAT	TAC	TTA	TTA	TCC	AAA	GGA	AGT	ATT	TTT	GAG	GCT	CTT	CCA	AAT	GAC	1661
156	S	D	A	M	H	Y	L	L	S	K	G	S	I	F	E	A	L	P	N	D	175
1662	AAT	TAC	CTT	CAG	ATT	TCT	GGC	ATA	CCA	CTT	TTT	AAA	AAT	AAT	GTG	TTT	GAG	GAA	ACT	GTG	1721
176	N	Y	L	Q	I	S	G	I	P	L	F	K	N	N	V	F	E	E	T	V	195
1722	TCA	AAA	AAA	AGA	AAG	CGA	ACC	ATT	GAA	ACA	TCC	ATT	ACT	CAA	AAT	AAA	AGC	GCC	CGC	AAA	1781
196	S	K	K	R	K	R	T	I	E	T	S	I	T	Q	N	K	S	A	R	K	215
1782	GAA	GTT	TCC	TGG	AAT	AGC	ATT	TCA	ATT	AGT	AGG	TTT	AGC	ATT	TTT	TAC	AGG	TCA	TCC	TAT	1841
216	E	V	S	W	N	S	I	S	I	S	R	F	S	I	F	Y	R	S	S	Y	235
1842	AAG	AAG	TTT	AAG	CAA	G	gtaacta	atactgtt	atccttc	ataacta	aattttag	AT	CTA	TAT	TTT	AAC					
1907																					
236	K	K	F	K	Q	D									L	Y	F	N		245	
1908	TTA	CAC	TCT	ATT	TGT	GAT	CGG	AAC	ACA	GTA	CAC	ATG	TGG	CTT	CAA	TGG	ATT	TTT	CCA	AGG	1967
246	L	H	S	I	C	D	R	N	T	V	H	M	W	L	Q	W	I	F	P	R	265
1968	CAA	TTT	GGA	CTT	ATA	AAC	GCA	TTT	CAA	GTG	AAG	CAA	TTG	CAC	AAA	GTG	ATT	CCA	CTG	GTA	2027
266	Q	F	G	L	I	N	A	F	Q	V	K	Q	L	H	K	V	I	P	L	V	285
2028	TCA	CAG	AGT	ACA	GTT	GTG	CCC	AAA	CGT	CTC	CTA	AAG	GTA	TAC	CCT	TTA	ATT	GAA	CAA	ACA	2087
286	S	Q	S	T	V	V	P	K	R	L	L	K	V	Y	P	L	I	E	Q	T	305
2088	GCA	AAG	CGA	CTC	CAT	CGT	ATT	TCT	CTA	TCA	AAA	GTT	TAC	AAC	CAT	TAT	TGC	CCA	TAT	ATT	2147
306	A	K	R	L	H	R	I	S	L	S	K	V	Y	N	H	Y	C	P	Y	I	325
2148	GAC	ACC	CAC	GAT	GAT	GAA	AAA	ATC	CTT	AGT	TAT	TCC	TTA	AAG	CCG	AAC	CAG	GTG	TTT	GCG	2207
326	D	T	H	D	D	E	K	I	L	S	Y	S	L	K	P	N	Q	V	F	A	345
2208	TTT	CTT	CGA	TCC	ATT	CTT	GTT	CGA	GTG	TTT	CCT	AAA	TTA	ATC	TGG	GGT	AAC	CAA	AGG	ATA	2267
346	F	L	R	S	I	L	V	R	V	F	P	K	L	I	W	G	N	Q	R	I	365
2268	TTT	GAG	ATA	ATA	TTA	AAA	G	gtattgt	ataaaa	atttatt	accacta	acgatttt	accag	AC	CTC	GAA	ACT	2336			
366	F	E	I	I	L	K	D							L	E	T		375			

FIGURE 46 (cont.)

2337	TTC	TTG	AAA	TTA	TCG	AGA	TAC	GAG	TCT	TTT	AGT	TTA	CAT	TAT	TTA	ATG	AGT	AAC	ATA	AAG	2396							
376	F	L	K	L	S	R	Y	E	S	F	S	L	H	Y	L	M	S	N	I	K	395							
2397	gtaatatgccaaatttttttaccattaattaacaatcag										ATT	TCA	GAA	ATT	GAA	TGG	CTA	GTC	CTT	GGA	2465							
396											I	S	E	I	E	W	L	V	L	G	405							
2466	AAA	AGG	TCA	AAT	GCG	AAA	ATG	TGC	TTA	AGT	GAT	TTT	GAG	AAA	CGC	AAG	CAA	ATA	TTT	GCG	2525							
406	K	R	S	N	A	K	M	C	L	S	D	F	E	K	R	K	Q	I	F	A	425							
2526	GAA	TTC	ATC	TAC	TGG	CTA	TAC	AAT	TCG	TTT	ATA	ATA	CCT	ATT	TTA	CAA	TCT	TTT	TTT	TAT	2585							
426	E	F	I	Y	W	L	Y	N	S	F	I	I	P	I	L	Q	S	F	F	Y	445							
2586	ATC	ACT	GAA	TCA	AGT	GAT	TTA	CGA	AAT	CGA	ACT	GTT	TAT	TTT	AGA	AAA	GAT	ATT	TGG	AAA	2645							
446	I	T	E	S	S	D	L	R	N	R	T	V	Y	F	R	K	D	I	W	K	465							
2646	CTC	TTG	TGC	CGA	CCC	TTT	ATT	ACA	TCA	ATG	AAA	ATG	GAA	GCG	TTT	GAA	AAA	ATA	AAC	GAG	2705							
466	L	L	C	R	P	F	I	T	S	M	K	M	E	A	F	E	K	I	N	E	485							
2706	gtatttttaaagtattttttgcaaaaagctaataatttttcag										AAC	AAT	GTT	AGG	ATG	GAT	ACT	CAG	AAA	ACT	2775							
486											N	N	V	R	M	D	T	Q	K	T	495							
2776	ACT	TTG	CCT	CCA	GCA	GTT	ATT	CGT	CTA	TTA	CCT	AAG	AAG	AAT	ACC	TTT	CGT	CTC	ATT	ACG	2835							
496	T	L	P	P	A	V	I	R	L	L	P	K	K	N	T	F	R	L	I	T	515							
2836	AAT	TTA	AGA	AAA	AGA	TTC	TTA	ATA	AAG	gtattaattttttggtcatcaatgtactttacttctaatactatta										2906								
516	N	L	R	K	R	F	L	I	K											524								
2907	ttagcag	ATG	GGT	TCA	AAC	AAA	AAA	ATG	TTA	GTC	AGT	ACG	AAC	CAA	ACT	TTA	CGA	CCT	GTG	2967								
525		M	G	S	N	K	K	M	L	V	S	T	N	Q	T	L	R	P	V	542								
2968	GCA	TCG	ATA	CTG	AAA	CAT	TTA	ATC	AAT	GAA	GAA	AGT	AGT	GGT	ATT	CCA	TTT	AAC	TTG	GAG	3027							
543	A	S	I	L	K	H	L	I	N	E	E	S	S	G	I	P	F	N	L	E	562							
3028	GTT	TAC	ATG	AAG	CTT	CTT	ACT	TTT	AAG	AAG	GAT	CTT	CTT	AAG	CAC	CGA	ATG	TTT	GG	gtaat	3088							
563	V	Y	M	K	L	L	T	F	K	K	D	L	L	K	H	R	M	F	G		581							
3089	tatataatgcgcgattcctcattattaattttgcag										G	CGT	AAG	AAG	TAT	TTT	GTA	CGG	ATA	GAT	ATA	3155						
582												R	K	K	Y	F	V	R	I	D	I	591						
3156	AAA	TCC	TGT	TAT	GAT	CGA	ATA	AAG	CAA	GAT	TTG	ATG	TTT	CGG	ATT	GTT	AAA	AAG	AAA	CTC	3215							
592	K	S	C	Y	D	R	I	K	Q	D	L	M	F	R	I	V	K	K	K	L	611							
3216	AAG	GAT	CCC	GAA	TTT	GTA	ATT	CGA	AAG	TAT	GCA	ACC	ATA	CAT	GCA	ACA	AGT	GAC	CGA	GCT	3275							
612	K	D	P	E	F	V	I	R	K	Y	A	T	I	H	A	T	S	D	R	A	631							
3276	ACA	AAA	AAC	TTT	GTT	AGT	GAG	GCG	TTT	TCC	TAT	T	gtaagtttattttttcattggaattttttaacaa							3343								
632	T	K	N	F	V	S	E	A	F	S	Y	F								643								
3344	attcttttttag										TT	GAT	ATG	GTG	CCT	TTT	GAA	AAA	GTC	GTG	CAG	TTA	CTT	TCT	ATG	AAA	ACA	3405
644												D	M	V	P	F	E	K	V	V	Q	L	L	S	M	K	T	659
3406	TCA	GAT	ACT	TTG	TTT	GTT	GAT	TTT	GTG	GAT	TAT	TGG	ACC	AAA	AGT	TCT	TCT	GAA	ATT	TTT	3465							
660	S	D	T	L	F	V	D	F	V	D	Y	W	T	K	S	S	S	E	I	F	679							
3466	AAA	ATG	CTC	AAG	GAA	CAT	CTC	TCT	GGA	CAC	ATT	GTT	AAG	gtataccaattgttgaattgtaataaca							3532							
680	K	M	L	K	E	H	L	S	G	H	I	V	K								692							

3533	cta	atg	aa	act	tag	ATA	GGA	AAT	TCT	CAA	TAC	CTT	CAA	AAA	GTT	GGT	ATC	CCT	CAG	GGC	TCA	3593
693						I	G	N	S	Q	Y	L	Q	K	V	G	I	P	Q	G	S	708
3594	ATT	CTG	TCA	TCT	TTT	TTG	TGT	CAT	TTC	TAT	ATG	GAA	GAT	TTG	ATT	GAT	GAA	TAC	CTA	TCG	3653	
709	I	L	S	S	F	L	C	H	F	Y	M	E	D	L	I	D	E	Y	L	S	728	
3654	TTT	ACG	AAA	AAG	AAA	GGA	TCA	GTG	TTG	TTA	CGA	GTA	GTC	GAC	GAT	TTC	CTC	TTT	ATA	ACA	3713	
729	F	T	K	K	K	G	S	V	L	L	R	V	V	D	D	F	L	F	I	T	748	
3714	GTT	AAT	AAA	AAG	GAT	GCA	AAA	AAA	TTT	TTG	AAT	TTA	TCT	TTA	AGA	G	gtgagattgctgtcattcc	3777				
749	V	N	K	K	D	A	K	K	F	L	N	L	S	L	R	G		764				
3778	taagttc	taaccg	ttgaag			GA	TTT	GAG	AAA	CAC	AAT	TTT	TCT	ACG	AGC	CTG	GAG	AAA	ACA	GTA	3840	
765						F	E	K	H	N	F	S	T	S	L	E	K	T	V		778	
3841	ATA	AAC	TTT	GAA	AAT	AGT	AAT	GGG	ATA	ATA	AAC	AAT	ACT	TTT	TTT	AAT	GAA	AGC	AAG	AAA	3900	
779	I	N	F	E	N	S	N	G	I	I	N	N	T	F	F	N	E	S	K	K	798	
3901	AGA	ATG	CCA	TTC	TTC	GGT	TTC	TCT	GTG	AAC	ATG	AGG	TCT	CTT	GAT	ACA	TTG	TTA	GCA	TGT	3960	
799	R	M	P	F	F	G	F	S	V	N	M	R	S	L	D	T	L	L	A	C	818	
3961	CCT	AAA	ATT	GAT	GAA	GCC	TTA	TTT	AAC	TCT	ACA	TCT	GTA	GAG	CTG	ACG	AAA	CAT	ATG	GGG	4020	
819	P	K	I	D	E	A	L	F	N	S	T	S	V	E	L	T	K	H	M	G	838	
4021	AAA	TCT	TTT	TTT	TAC	AAA	ATT	CTA	AG	gtatactgtgtaactgaataatagctgacaaataatcag	A	TCG	4089									
839	K	S	F	F	Y	K	I	L	R		S		848									
4090	AGC	CTT	GCA	TCC	TTT	GCA	CAA	GTA	TTT	ATT	GAC	ATT	ACC	CAC	AAT	TCA	AAA	TTC	AAT	TCT	4149	
849	S	L	A	S	F	A	Q	V	F	I	D	I	T	H	N	S	K	F	N	S	868	
4150	TGC	TGC	AAT	ATA	TAT	AGG	CTA	GGA	TAC	TCT	ATG	TGT	ATG	AGA	GCA	CAA	GCA	TAC	TTA	AAA	4209	
869	C	C	N	I	Y	R	L	G	Y	S	M	C	M	R	A	Q	A	Y	L	K	888	
4210	AGG	ATG	AAG	GAT	ATA	TTT	ATT	CCC	CAA	AGA	ATG	TTC	ATA	ACG	G	gtgagtacttatttttaactaga	4274					
889	R	M	K	D	I	F	I	P	Q	R	M	F	I	T	D		903					
4275	aaagtcattaattaac	ccttag				AT	CTT	TTG	AAT	GTT	ATT	GGA	AGA	AAA	ATT	TGG	AAA	AAG	TTG	GCC	4339	
904						L	L	N	V	I	G	R	K	I	W	K	K	L	A		917	
4340	GAA	ATA	TTA	GGA	TAT	ACG	AGT	AGG	CGT	TTC	TTG	TCC	TCT	GCA	GAA	GTC	AAA	TG	gtacgtgtc	4401		
918	E	I	L	G	Y	T	S	R	R	F	L	S	S	A	E	V	K	W		935		
4402	ggtctcgagacttcag	caatattgacacatcag	G			CTT	TTT	TGT	CTT	GGA	ATG	AGA	GAT	GGT	TTG	AAA	4468					
936						L	F	C	L	G	M	R	D	G	L	K	946					
4469	CCC	TCT	TTC	AAA	TAT	CAT	CCA	T														

100

1. **NAME** _____
 2. **ADDRESS** _____
 3. **CITY** _____
 4. **STATE** _____
 5. **ZIP** _____
 6. **PHONE** _____
 7. **DATE** _____
 8. **SIGNATURE** _____
 9. **PRINT NAME** _____
 10. **PRINT ADDRESS** _____
 11. **PRINT CITY** _____
 12. **PRINT STATE** _____
 13. **PRINT ZIP** _____
 14. **PRINT PHONE** _____
 15. **PRINT DATE** _____
 16. **PRINT SIGNATURE** _____
 17. **PRINT NAME** _____
 18. **PRINT ADDRESS** _____
 19. **PRINT CITY** _____
 20. **PRINT STATE** _____
 21. **PRINT ZIP** _____
 22. **PRINT PHONE** _____
 23. **PRINT DATE** _____
 24. **PRINT SIGNATURE** _____
 25. **PRINT NAME** _____
 26. **PRINT ADDRESS** _____
 27. **PRINT CITY** _____
 28. **PRINT STATE** _____
 29. **PRINT ZIP** _____
 30. **PRINT PHONE** _____
 31. **PRINT DATE** _____
 32. **PRINT SIGNATURE** _____
 33. **PRINT NAME** _____
 34. **PRINT ADDRESS** _____
 35. **PRINT CITY** _____
 36. **PRINT STATE** _____
 37. **PRINT ZIP** _____
 38. **PRINT PHONE** _____
 39. **PRINT DATE** _____
 40. **PRINT SIGNATURE** _____
 41. **PRINT NAME** _____
 42. **PRINT ADDRESS** _____
 43. **PRINT CITY** _____
 44. **PRINT STATE** _____
 45. **PRINT ZIP** _____
 46. **PRINT PHONE** _____
 47. **PRINT DATE** _____
 48. **PRINT SIGNATURE** _____
 49. **PRINT NAME** _____
 50. **PRINT ADDRESS** _____
 51. **PRINT CITY** _____
 52. **PRINT STATE** _____
 53. **PRINT ZIP** _____
 54. **PRINT PHONE** _____
 55. **PRINT DATE** _____
 56. **PRINT SIGNATURE** _____
 57. **PRINT NAME** _____
 58. **PRINT ADDRESS** _____
 59. **PRINT CITY** _____
 60. **PRINT STATE** _____
 61. **PRINT ZIP** _____
 62. **PRINT PHONE** _____
 63. **PRINT DATE** _____
 64. **PRINT SIGNATURE** _____
 65. **PRINT NAME** _____
 66. **PRINT ADDRESS** _____
 67. **PRINT CITY** _____
 68. **PRINT STATE** _____
 69. **PRINT ZIP** _____
 70. **PRINT PHONE** _____
 71. **PRINT DATE** _____
 72. **PRINT SIGNATURE** _____
 73. **PRINT NAME** _____
 74. **PRINT ADDRESS** _____
 75. **PRINT CITY** _____
 76. **PRINT STATE** _____
 77. **PRINT ZIP** _____
 78. **PRINT PHONE** _____
 79. **PRINT DATE** _____
 80. **PRINT SIGNATURE** _____
 81. **PRINT NAME** _____
 82. **PRINT ADDRESS** _____
 83. **PRINT CITY** _____
 84. **PRINT STATE** _____
 85. **PRINT ZIP** _____
 86. **PRINT PHONE** _____
 87. **PRINT DATE** _____
 88. **PRINT SIGNATURE** _____
 89. **PRINT NAME** _____
 90. **PRINT ADDRESS** _____
 91. **PRINT CITY** _____
 92. **PRINT STATE** _____
 93. **PRINT ZIP** _____
 94. **PRINT PHONE** _____
 95. **PRINT DATE** _____
 96. **PRINT SIGNATURE** _____
 97. **PRINT NAME** _____
 98. **PRINT ADDRESS** _____
 99. **PRINT CITY** _____
 100. **PRINT STATE** _____
 101. **PRINT ZIP** _____
 102. **PRINT PHONE** _____
 103. **PRINT DATE** _____
 104. **PRINT SIGNATURE** _____
 105. **PRINT NAME** _____
 106. **PRINT ADDRESS** _____
 107. **PRINT CITY** _____
 108. **PRINT STATE** _____
 109. **PRINT ZIP** _____
 110. **PRINT PHONE** _____
 111. **PRINT DATE** _____
 112. **PRINT SIGNATURE** _____
 113. **PRINT NAME** _____
 114. **PRINT ADDRESS** _____
 115. **PRINT CITY** _____
 116. **PRINT STATE** _____
 117. **PRINT ZIP** _____
 118. **PRINT PHONE** _____
 119. **PRINT DATE** _____
 120. **PRINT SIGNATURE** _____
 121. **PRINT NAME** _____
 122. **PRINT ADDRESS** _____
 123. **PRINT CITY** _____
 124. **PRINT STATE** _____
 125. **PRINT ZIP** _____
 126. **PRINT PHONE** _____
 127. **PRINT DATE** _____
 128. **PRINT SIGNATURE** _____
 129. **PRINT NAME** _____
 130. **PRINT ADDRESS** _____
 131. **PRINT CITY** _____
 132. **PRINT STATE** _____
 133. **PRINT ZIP** _____
 134. **PRINT PHONE** _____
 135. **PRINT DATE** _____
 136. **PRINT SIGNATURE** _____
 137. **PRINT NAME** _____
 138. **PRINT ADDRESS** _____
 139. **PRINT CITY** _____
 140. **PRINT STATE** _____
 141. **PRINT ZIP** _____
 142. **PRINT PHONE** _____
 143. **PRINT DATE** _____
 144. **PRINT SIGNATURE** _____
 145. **PRINT NAME** _____
 146. **PRINT ADDRESS** _____
 147. **PRINT CITY** _____
 148. **PRINT STATE** _____
 149. **PRINT ZIP** _____
 150. **PRINT PHONE** _____
 151. **PRINT DATE** _____
 152. **PRINT SIGNATURE** _____
 153. **PRINT NAME** _____
 154. **PRINT ADDRESS** _____
 155. **PRINT CITY** _____
 156. **PRINT STATE** _____
 157. **PRINT ZIP** _____
 158. **PRINT PHONE** _____
 159. **PRINT DATE** _____
 160. **PRINT SIGNATURE** _____
 161. **PRINT NAME** _____
 162. **PRINT ADDRESS** _____
 163. **PRINT CITY** _____
 164. **PRINT STATE** _____
 165. **PRINT ZIP** _____
 166. **PRINT PHONE** _____
 167. **PRINT DATE** _____
 168. **PRINT SIGNATURE** _____
 169. **PRINT NAME** _____
 170. **PRINT ADDRESS** _____
 171. **PRINT CITY** _____
 172. **PRINT STATE** _____
 173. **PRINT ZIP** _____
 174. **PRINT PHONE** _____
 175. **PRINT DATE** _____
 176. **PRINT SIGNATURE** _____
 177. **PRINT NAME** _____
 178. **PRINT ADDRESS** _____
 179. **PRINT CITY** _____
 180. **PRINT STATE** _____
 181. **PRINT ZIP** _____
 182. **PRINT PHONE** _____
 183. **PRINT DATE** _____
 184. **PRINT SIGNATURE** _____
 185. **PRINT NAME** _____
 186. **PRINT ADDRESS** _____
 187. **PRINT CITY** _____
 188. **PRINT STATE** _____
 189. **PRINT ZIP** _____
 190. **PRINT PHONE** _____
 191. **PRINT DATE** _____
 192. **PRINT SIGNATURE** _____
 193. **PRINT NAME** _____
 194. **PRINT ADDRESS** _____
 195. **PRINT CITY** _____
 196. **PRINT STATE** _____
 197. **PRINT ZIP** _____
 198. **PRINT PHONE** _____
 199. **PRINT DATE** _____
 200. **PRINT SIGNATURE** _____
 201. **PRINT NAME** _____
 202. **PRINT ADDRESS** _____
 203. **PRINT CITY** _____
 204. **PRINT STATE** _____
 205. **PRINT ZIP** _____
 206. **PRINT PHONE** _____
 207. **PRINT DATE** _____
 208. **PRINT SIGNATURE** _____
 209. **PRINT NAME** _____
 210. **PRINT ADDRESS** _____
 211. **PRINT CITY** _____
 212. **PRINT STATE** _____
 213. **PRINT ZIP** _____
 214. **PRINT PHONE** _____
 215. **PRINT DATE** _____
 216. **PRINT SIGNATURE** _____
 217. **PRINT NAME** _____
 218. **PRINT ADDRESS** _____
 219. **PRINT CITY** _____
 220. **PRINT STATE** _____

FIGURE 47

[illegible]

FIGURE 47 (cont.)

190	thr	tyr	cys	val	arg	arg	tyr	ala	val	val	200	gln	lys	ala	ala	met
	ACG	TAC	TGC	GTG	CGT	CGG	TAT	GCC	GTG	GTC		CAG	AAG	GCC	GCC	ATG
											210	gly	thr	ser	ala	arg
	gly	thr	ser	ala	arg	pro	ser	arg	ala	thr		ser	tyr	val	gln	cys
	GGC	ACG	TCC	GCA	AGG	CCT	TCA	AGA	GCC	ACG		TCC	TAC	GTC	CAG	TGC
220	gln	gly	ile	pro	gln	gly	ser	ile	leu	ser	230	thr	leu	leu	cys	ser
	CAG	GGG	ATC	CCG	CAG	GGC	TCC	ATC	CTC	TCC		ACG	CTG	CTC	TGC	AGC
											240	leu	cys	tyr	gly	asp
	leu	cys	tyr	gly	asp	met	glu	asn	lys	leu		phe	ala	gly	ile	arg
	CTG	TGC	TAC	GGC	GAC	ATG	GAG	AAC	AAG	CTG		TTT	GCG	GGG	ATT	CGG
250	arg	asp	gly	leu	leu	leu	arg	leu	val	asp	260	asp	phe	leu	leu	val
	CGG	GAC	GGG	CTG	CTC	CTG	CGT	TTG	GTG	GAT		GAT	TTC	TTG	TTG	GTG
											270	thr	pro	his	leu	thr
	thr	pro	his	leu	thr	his	ala	lys	thr	phe		leu	arg	thr	leu	val
	ACA	CCT	CAC	CTC	ACC	CAC	GCG	AAA	ACC	TTC		CTC	AGG	ACC	CTG	GTC
280	arg	gly	val	pro	glu	tyr	gly	cys	val	val	290	asn	leu	arg	lys	thr
	CGA	GGT	GTC	CCT	GAG	TAT	GGC	TGC	GTG	GTG		AAC	TTG	CGG	AAG	ACA
											300	val	val	asn	phe	pro
	val	val	asn	phe	pro	val	glu	asp	glu	ala		leu	gly	gly	thr	ala
	GTG	GTG	AAC	TTC	CCT	GTA	GAA	GAC	GAG	GCC		CTG	GGT	GGC	ACG	GCT
310	phe	val	gln	met	pro	ala	his	gly	leu	phe	320	pro	trp	cys	gly	leu
	TTT	GTT	CAG	ATG	CCG	GCC	CAC	GGC	CTA	TTC		CCC	TGG	TGC	GGC	CTG
											330	leu	leu	asp	thr	arg
	leu	leu	asp	thr	arg	thr	leu	glu	val	gln		ser	asp	tyr	ser	ser
	CTG	CTG	GAT	ACC	CGG	ACC	CTG	GAG	GTG	CAG		AGC	GAC	TAC	TCC	AGC
340	tyr	ala	arg	thr	ser	ile	arg	ala	ser	leu	350	thr	phe	asn	arg	gly
	TAT	GCC	CGG	ACC	TCC	ATC	AGA	GCC	AGT	CTC		ACC	TTC	AAC	CGC	GGC
											360	phe	lys	ala	gly	arg
	phe	lys	ala	gly	arg	asn	met	arg	arg	lys		leu	phe	gly	val	leu
	TTC	AAG	GCT	GGG	AGG	AAC	ATG	CGT	CGC	AAA		CTC	TTT	GGG	GTC	TTG
370	arg	leu	lys	cys	his	ser	leu	phe	leu	asp	380	leu	gln	val	asn	ser
	CGG	CTG	AAG	TGT	CAC	AGC	CTG	TTT	CTG	GAT		TTG	CAG	GTG	AAC	AGC

THE
NEW
YORK
PUBLIC
LIBRARY

390
leu gln thr val cys thr asn ile tyr lys ile leu leu leu gln
CTC CAG ACG GTG TGC ACC AAC ATC TAC AAG ATC CTC CTG CTG CAG

400
ala tyr arg phe his ala cys val leu gln leu pro phe his gln
GCG TAC AGG TTT CAC GCA TGT GTG CTG CAG CTC CCA TTT CAT CAG

420
gln val trp lys asn pro his phe ser cys ala ser ser leu thr
CAA GTT TGG AAG AAC CCA CAT TTT TCC TGC GCG TCA TCT CTG ACA

430
arg leu pro leu leu leu his pro glu ser gln glu arg arg asp
CGG CTC CCT CTG CTA CTC CAT CCT GAA AGC CAA GAA CGC AGG GAT

450
val ala gly gly gln gly arg arg arg pro ser ala leu arg gly
GTC GCT GGG GGC CAA GGG CGC CGC CGG CCC TCT GCC CTC CGA GGC

460
arg ala val ala val pro pro ser ile pro ala gln ala asp ser
CGT GCA GTG GCT GTG CCA CCA AGC ATT CCT GCT CAA GCT GAC TCG

480
thr pro cys his leu arg ala thr pro gly val thr gln asp ser
ACA CCG TGT CAC CTA CGT GCC ACT CCT GGG GTC ACT CAG GAC AGC

490
pro asp ala ala glu ser glu ala pro gly asp asp ala asp cys
CCA GAC GCA GCT GAG TCG GAA GCT CCC GGG GAC GAC GCT GAC TGC

510
pro gly gly arg ser gln pro gly thr ala leu arg leu gln asp
CCT GGA GGC CGC AGC CAA CCC GGC ACT GCC CTC AGA CTT CAA GAC

520
his pro gly leu met ala thr arg pro gln pro gly arg glu gln
CAT CCT GGA CTG ATG GCC ACC CGC CCA CAG CCA GGC CGA GAG CAG

540
thr pro ala ala leu ser arg arg ala tyr thr ser gln gly gly
ACA CCA GCA GCC CTG TCA CGC CGG GCT TAT ACG TCC CAG GGA GGG

550
arg gly gly pro his pro gly leu his arg trp glu ser glu ala
AGG GGC GGC CCA CAC CCA GGC CTG CAC CGC TGG GAG TCT GAG GCC

560
arg
OP
TGA GTGAGTGT TTTGGCCGAGGCCTGCATGTCCGGCTGAAGGCTGAGTGTCCGGCTGAGGC
CTGAGCGAGTGTCCAGCCAAGGGCTGAGTGTCCAGCACACCTGCGTTTTCACTTCCCCAC

8

[illegible]

FIGURE 48

Motif -1	
Ep p123	...LVVSLIRCFFYVTEQQKSYSKT...
Sp Tez1	...FIIPILQSFFYITESDLRNRT...
Sc Est2	...LIPKIIQTFFYCTEISSTVTIV...
Hs TCP1	...YVVELLRSFFYVTETTFQKNRL...
consensus	FFY TE
Motif 0	K
	p hhh K hR h R
Ep p123	...KSLGFAPGKLRRLIPKKT--TFRPIMTFNKKIV....
Sp Tez1	...QKTTLPPAVIRLLPKKN--TFRLITNLRKRFL...
Sc Est2	...TLSNFNHSMRIIPKKSNEFRJIAIPCRGAD...
Hs TCP1	...ARPALLTSRLRFIPKPD--GLRPVNMMDYVVG...
consensus	R PK R I
Motif A	AF
	h hDh GY h
Ep p123	...PKLFFATMDIEKCYDSVNREKLSTFLK...
Sp Tez1	...RKKYFVRIDIKSCYDRIKQDLMFRIVK...
Sc Est2	...PELYFMKFDVKSCYDSIPRMECMRILK...
Hs TCP1	...PELYFVKVDVTGAYDTIPQDRLTEVIA...//...
consensus	F D YD
Motif B	hPQG pS hh
Ep p123	...NGKFYKQTKGIPQGLCVSSILSSFYYA...
Sp Tez1	...GNSQYLQKVGIPQGSILSSFLCHFYME...
Sc Est2	...EDKCYIREDGLFQGSSLSAPIVDLVYD...
Hs TCP1	...RATSYVQCQGIPQGSILSTLLCSLCYG...
consensus	G QG S
Motif C	Y
	h F DDhhh
Ep p123	...PNVNLLMRLTDDYLLITTQENN...
Sp Tez1	...KKGSVLLRVVDDFLFITVNKKD...
Sc Est2	...SQDTLILKLADDFLIISTDQQQ...
Hs TCP1	...RRDGLLLRLVDDFLLVTPHLTH...
consensus	DD L
Motif D	Gh h cK
Ep p123	...NVSRENGFKFNMKKL...
Sp Tez1	...LNLSLRGFEEKHNFST...
Sc Est2	...KKLAMGGFQKYNAKA...
Hs TCP1	...LRTLVRGVPEYGCVV...
consensus	G

FIGURE 49

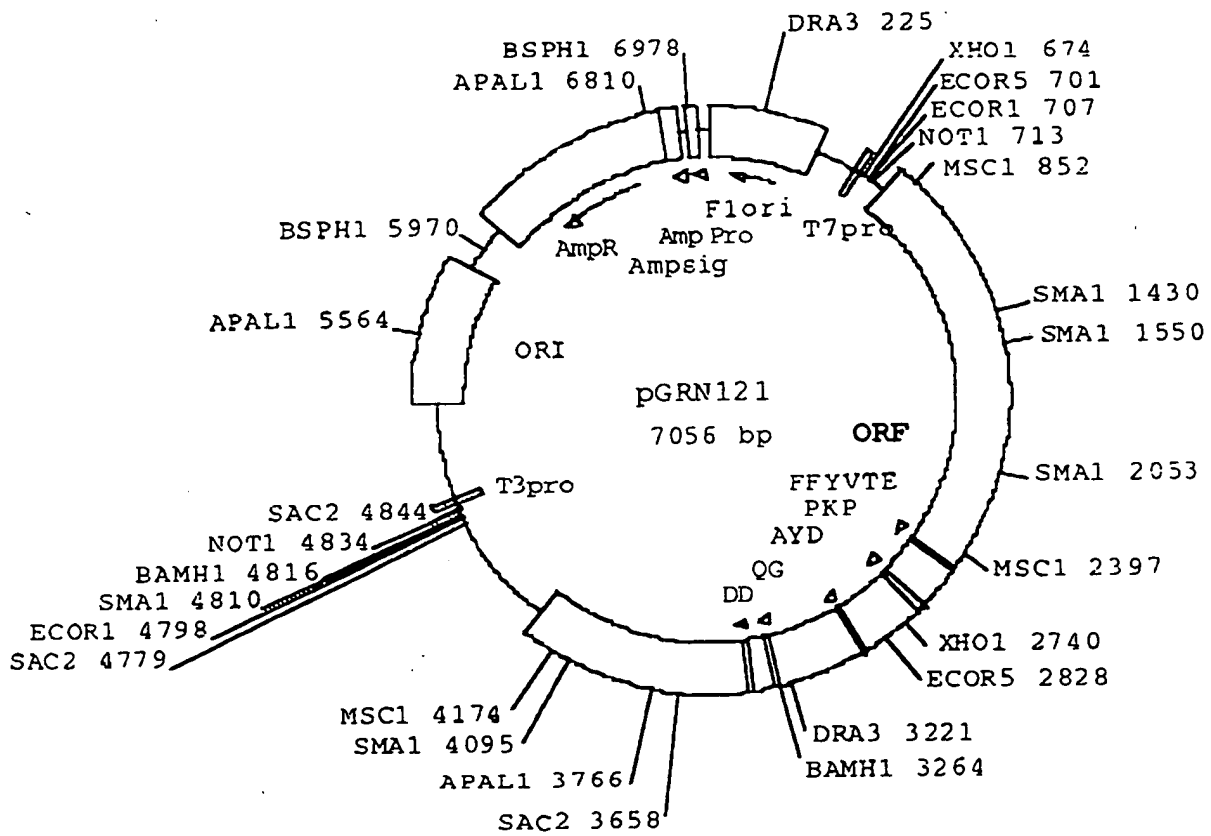


FIGURE 50

1 GCAGCGCTGC GTCCTGCTGC GCACGTGGGA AGCCCTGGCC CCGGCCACCC
 51 CCGCGATGCC GCGCGCTCCC CGCTGCCGAG CCGTGCGCTC CCTGCTGCGC
 101 AGCCACTACC GCGAGGTGCT GCCGCTGGCC ACGTTCGTGC GCGCCTGGG
 151 GCCCCAGGGC TGGCGGCTGG TGCAGCGCGG GGACCCGGCG GCTTCCGCG
 201 CGNTGGTGGC CCANTGCNTG GTGTGCGTGC CCTGGGANGN ANGGCNGCCC
 251 CCCGCCGCC CCTCCTTCCG CCAGGTGTCC TGCCTGAANG ANCTGGTGGC
 301 CCGAGTGCTG CANANGCTGT GCGANCGCGG CGCGAANAAC GTGCTGGCCT
 351 TCGGCTTCGC GCTGCTGGAC GGGGCCCGCG GGGGCCCCC CGAGGCCTTC
 401 ACCACCAGCG TGCAGCTA CCTGCCAAC ACGGTGACCG ACGCACTGCG
 451 GGGGAGCGGG GCGTGGGGG TGCTGCTGCG CCGCGTGGGC GACGACGTGC
 501 TGGTTCACCT GCTGGCACGC TGC GCGTNT TTGTGCTGGT GGNTCCAGC
 551 TGC GCTACC ANGTGTGCGG GCCGCCGCTG TACCAGCTCG GCGCTGCNAC
 601 TCAGGCCCGG CCCCCGCCAC ACGTANTGG ACCCGAANGC GTCTGGGATC
 651 CAACGGGCCT GGAACCATAG CGTCAGGGAG GCCGGGGTCC CCCTGGGCTG
 701 CCAGCCCCGG GTGCGAGGAG GCGCGGGGGC AGTGCCAGCC GAAGTCTGCC
 751 GTTGCCCAAG AGGCCAGGC GTGGCGCTGC CCCTGAGCCG GAGCGGACGC
 801 CCGTTGGGCA GGGTCTTG GCCACCCGG GCAGGACGCC TGGACCGAGT
 851 GACCGTGGTT TCTGTGTGGT GTCACCTGCC AGACCCGCCG AAGAAGCCAC
 901 CTCTTTGGAG GGTGCGCTCT CTGGCACGCG CCACTCCAC CCATCCGTGG
 951 GCCGCCAGCA CCACGCGGGC CCCCATCCA CATCGCGGCC ACCACGTCCT
 1001 GGGACACGCC TTGTCCCCCG GTGTACGCC AGACCAAGCA CTTCTCTAC
 1051 TCCTCAGGCG ACAAGNACAC TGCGNCCCTC CTTCTACTC AATATATCTG
 1101 AGGCCAGCC TGA CTGGCGT TCGGGAGGTT CGTGGAGACA NTCTTTCTGG
 1151 TTCCAGGCCT TGGATGCCAG GATTCCCCGC AGGTTGCCCC GCCTGCCCCA
 1201 GCGN TACTGG CAAATGCGGC CCCTGTTTCT GGAGCTGCTT GGAACCACG
 1251 CGCAGTGCCC CTACGGGGTG TTCCTCAAGA CGCACTGCCC GCTGCGAGCT
 1301 GCGGTCACCC CAGCAGCCGG TGTCTGTGCC CGGGAGAAGC CCCAGGGCTC
 1351 TGTGGCGGCC CCCGAGGAGG AGGAACACAG ACCCCGTCG CCTGGTGCAG
 1401 CTGCTCCGCC AGCACAGCAG CCCCTGGCAG GTGTACGGCT TCGTGCGGGC
 1451 CTGCCTGCGC CGGCTGGTGC CCCCAGGCCT CTGGGGCTCC AGGCACAACG
 1501 AACGCCGCTT CCTCAGGAAC ACCAAGAAGT TCATCTCCCT GGGGAAGCAT
 1551 GCCAAGCTCT CGCTGCAGGA GCTGACGTGG AAGATGAGCG TGCGGGACTG
 1601 CGCTTGGCTG CGCAGGAGCC CAGGGGTTGG CTGTGTTCCG GCCGCAGAGC
 1651 ACCGTCTGCG TGAGGAGATC CTGGCCAAGT TCCTGCACTG GCTGATGAGT
 1701 GTGTACGTCG TCGAGCTGCT CAGGTCTTTC TTTTATGTCA CGGAGACCAC
 1751 GTTTCAAAAG AACAGGCTCT TTTTCTACCG GAAGAGTGTC TGGAGCAAGT
 1801 TGCAAAGCAT TGGAATCAGA CAGCACTTGA AGAGGGTGCA GCTGCGGGAG
 1851 CTGTGGAAG CAGAGGTCAG GCAGCATCGG GAAGCCAGGC CCGCCCTGCT
 1901 GACGTCCAGA CTCCGCTTCA TCCCAAGCC TGACGGGCTG CGGCCGATTG
 1951 TGAACATGGA CTACGTCGTG GGAGCCAGAA CGTCCGCAG AGAAAAGAGG
 2001 GCCGAGCGTC TCACCTCGAG GGTGAAGGCA CTGTTCAGCG TGCTCAACTA
 2051 CGAGCGGGCG CGGCGCCCCG GCCTCCTGGG CGCCTCTGTG CTGGGCCTGG
 2101 ACGATATCCA CAGGGCCTGG CGCACCTTCG TGCTGCGTGT GCGGGCCAG
 2151 GACCCGCCGC CTGAGCTGTA CTTTGTCAAG GTGGATGTGA CGGGCGCGTA
 2201 CGACACCATC CCCCAGGACA GGCTCACGGA GGTCATCGCC AGCATCATCA
 2251 AACCCAGAA CACGTACTGC GTGCGTCGGT ATGCCGTGGT CCAGAAGGCC

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100

FIGURE 50 (cont.)

2301 GCCCATGGGC ACGTCCGCAA GGCCTTCAAG AGCCACGTCT CTACCTTGAC
 2351 AGACCTCCAG CCGTACATGC GACAGTTCGT GGCTCACCTG CAGGANAACA
 2401 GCCCGCTGAG GGATGCCGTC GTCATCGAGC AGAGCTCCTC CCTGAATGAG
 2451 GCCAGCAGTG GCCTCTTCGA CGTCTTCCTA CGCTTCATGT GCCACCACGC
 2501 CGTGCGCATC AGGGGCAAGT CCTACGTCCA GTGCCAGGGG ATCCCGCAGG
 2551 GCTCCATCCT CTCCACGCTG CTCTGCAGCC TGTGCTACGG CGACATGGAG
 2601 AACAAGCTGT TTGCGGGGAT TCGGCGGGAC GGGCTGCTCC TGC GTTTGGT
 2651 GGATGATTTT TTGTTGGTGA CACCTCACCT CACCCACGCG AAAACCTTCC
 2701 TCAGGACCCT GTCCGAGGT GTCCCTGAGT ATGGCTGCGT GGTGAACCTG
 2751 CGGAAGACAG TGGTGAACCT CCCTGTAGAA GACGAGGCC TGGGTGGCAC
 2801 GGCTTTTGT CAGATGCCGG CCCACGGCCT ATTCCCTGG TCGGGCCTGC
 2851 TGCTGGATAC CCGGACCCTG GAGGTGCAGA GCGACTACTC CAGCTATGCC
 2901 CGGACCTCCA TCAGAGCCAG TCTCACCTTC AACCGCGGCT TCAAGGCTGG
 2951 GAGGAACATG CGTCGCAAAC TCTTTGGGGT CTTGCGGCTG AAGTGTGACA
 3001 GCCTGTTTCT GGATTTGAG GTGAACAGCC TCCAGACGGT GTGCACCAAC
 3051 ATCTACAAGA TCCTCCTGCT GCAGGCGTAC AGGTTTCACG CATGTGTGCT
 3101 GCAGTCCCA TTTCATCAGC AAGTTTGAA GAACCCACACA TTTTTCCTGC
 3151 GCGTCATCTC TGACACGGCC TCCCTCTGCT ACTCCATCCT GAAAGCCAAG
 3201 AACGCAGGGA TGTCGCTGGG GGCCAAGGGC GCCGCCGGCC CTCTGCCCTC
 3251 CGAGGCCGTG CAGTGGCTGT GCCACCAAGC ATTCCTGCTC AAGCTGACTC
 3301 GACACCGTGT CACCTACGTG CCACTCCTGG GGTCACCTAG GACAGCCAG
 3351 ACGCAGCTGA GTCGGAAGCT CCCGGGGACG ACGCTGACTG CCCTGGAGGC
 3401 CGCAGCCAAC CCGGCACTGC CCTCAGACTT CAAGACCATC CTGGACTGAT
 3451 GGCCACCCGC CCACAGCCAG GCCGAGAGCA GACACCAGCA GCCCTGTCAC
 3501 GCCGGGCTCT ACGTCCCAGG GAGGGAGGGG CGGCCACAC CCAGGCCCGC
 3551 ACCGCTGGGA GTCTGAGGCC TGAGTGAGTG TTTGGCCGAG GCCTGCATGT
 3601 CCGGCTGAAG GCTGAGTGTC CGGCTGAGGC CTGAGCGAGT GTCCAGCCAA
 3651 GGGCTGAGTG TCCAGCACAC CTGCCGTCTT CACTTCCCCA CAGGCTGGCG
 3701 CTCGGCTCCA CCCCAGGGCC AGCTTTTCCT CACCAGGAGC CCGGCTTCCA
 3751 CTCCCACAT AGGAATAGTC CATCCCCAGA TTCGCCATTG TTCACCCCTC
 3801 GCCCTGCCCT CTTTGCCTT CCACCCCAAC CATCCAGGTG GAGACCCTGA
 3851 GAAGGACCCT GGGAGCTCTG GGAATTTGGA GTGACCAAAG GTGTGCCCTG
 3901 TACACAGGCG AGGACCCTGC ACCTGGATGG GGGTCCCTGT GGGTCAAATT
 3951 GGGGGGAGGT GCTGTGGGAG TAAATACTG AATATATGAG TTTTTCAGTT
 4001 TTGAAAAAAA AAAAAAAAAA AAAAAAAAAA

FIGURE 51

THE **NEW** **YORK** **PUBLIC** **LIBRARY**

[illegible][illegible]

1. NAME _____
 2. ADDRESS _____
 3. CITY _____
 4. STATE _____
 5. ZIP _____
 6. PHONE _____
 7. DATE _____
 8. SIGNATURE _____
 9. PRINT NAME _____
 10. DATE _____
 11. SIGNATURE _____
 12. PRINT NAME _____
 13. DATE _____
 14. SIGNATURE _____
 15. PRINT NAME _____
 16. DATE _____
 17. SIGNATURE _____
 18. PRINT NAME _____
 19. DATE _____
 20. SIGNATURE _____
 21. PRINT NAME _____
 22. DATE _____
 23. SIGNATURE _____
 24. PRINT NAME _____
 25. DATE _____
 26. SIGNATURE _____
 27. PRINT NAME _____
 28. DATE _____
 29. SIGNATURE _____
 30. PRINT NAME _____
 31. DATE _____
 32. SIGNATURE _____
 33. PRINT NAME _____
 34. DATE _____
 35. SIGNATURE _____
 36. PRINT NAME _____
 37. DATE _____
 38. SIGNATURE _____
 39. PRINT NAME _____
 40. DATE _____
 41. SIGNATURE _____
 42. PRINT NAME _____
 43. DATE _____
 44. SIGNATURE _____
 45. PRINT NAME _____
 46. DATE _____
 47. SIGNATURE _____
 48. PRINT NAME _____
 49. DATE _____
 50. SIGNATURE _____
 51. PRINT NAME _____
 52. DATE _____
 53. SIGNATURE _____
 54. PRINT NAME _____
 55. DATE _____
 56. SIGNATURE _____
 57. PRINT NAME _____
 58. DATE _____
 59. SIGNATURE _____
 60. PRINT NAME _____
 61. DATE _____
 62. SIGNATURE _____
 63. PRINT NAME _____
 64. DATE _____
 65. SIGNATURE _____
 66. PRINT NAME _____
 67. DATE _____
 68. SIGNATURE _____
 69. PRINT NAME _____
 70. DATE _____
 71. SIGNATURE _____
 72. PRINT NAME _____
 73. DATE _____
 74. SIGNATURE _____
 75. PRINT NAME _____
 76. DATE _____
 77. SIGNATURE _____
 78. PRINT NAME _____
 79. DATE _____
 80. SIGNATURE _____
 81. PRINT NAME _____
 82. DATE _____
 83. SIGNATURE _____
 84. PRINT NAME _____
 85. DATE _____
 86. SIGNATURE _____
 87. PRINT NAME _____
 88. DATE _____
 89. SIGNATURE _____
 90. PRINT NAME _____
 91. DATE _____
 92. SIGNATURE _____
 93. PRINT NAME _____
 94. DATE _____
 95. SIGNATURE _____
 96. PRINT NAME _____
 97. DATE _____
 98. SIGNATURE _____
 99. PRINT NAME _____
 100. DATE _____
 101. SIGNATURE _____
 102. PRINT NAME _____
 103. DATE _____
 104. SIGNATURE _____
 105. PRINT NAME _____
 106. DATE _____
 107. SIGNATURE _____
 108. PRINT NAME _____
 109. DATE _____
 110. SIGNATURE _____
 111. PRINT NAME _____
 112. DATE _____
 113. SIGNATURE _____
 114. PRINT NAME _____
 115. DATE _____
 116. SIGNATURE _____
 117. PRINT NAME _____
 118. DATE _____
 119. SIGNATURE _____
 120. PRINT NAME _____
 121. DATE _____
 122. SIGNATURE _____
 123. PRINT NAME _____
 124. DATE _____
 125. SIGNATURE _____
 126. PRINT NAME _____
 127. DATE _____
 128. SIGNATURE _____
 129. PRINT NAME _____
 130. DATE _____
 131. SIGNATURE _____
 132. PRINT NAME _____
 133. DATE _____
 134. SIGNATURE _____
 135. PRINT NAME _____
 136. DATE _____
 137. SIGNATURE _____
 138. PRINT NAME _____
 139. DATE _____
 140. SIGNATURE _____
 141. PRINT NAME _____
 142. DATE _____
 143. SIGNATURE _____
 144. PRINT NAME _____
 145. DATE _____
 146. SIGNATURE _____
 147. PRINT NAME _____
 148. DATE _____
 149. SIGNATURE _____
 150. PRINT NAME _____
 151. DATE _____
 152. SIGNATURE _____
 153. PRINT NAME _____
 154. DATE _____
 155. SIGNATURE _____
 156. PRINT NAME _____
 157. DATE _____
 158. SIGNATURE _____
 159. PRINT NAME _____
 160. DATE _____
 161. SIGNATURE _____
 162. PRINT NAME _____
 163. DATE _____
 164. SIGNATURE _____
 165. PRINT NAME _____
 166. DATE _____
 167. SIGNATURE _____
 168. PRINT NAME _____
 169. DATE _____
 170. SIGNATURE _____
 171. PRINT NAME _____
 172. DATE _____
 173. SIGNATURE _____
 174. PRINT NAME _____
 175. DATE _____
 176. SIGNATURE _____
 177. PRINT NAME _____
 178. DATE _____
 179. SIGNATURE _____
 180. PRINT NAME _____
 181. DATE _____
 182. SIGNATURE _____
 183. PRINT NAME _____
 184. DATE _____
 185. SIGNATURE _____
 186. PRINT NAME _____
 187. DATE _____
 188. SIGNATURE _____
 189. PRINT NAME _____
 190. DATE _____
 191. SIGNATURE _____
 192. PRINT NAME _____
 193. DATE _____
 194. SIGNATURE _____
 195. PRINT NAME _____
 196. DATE _____
 197. SIGNATURE _____
 198. PRINT NAME _____
 199. DATE _____
 200. SIGNATURE _____
 201. PRINT NAME _____
 202. DATE _____
 203. SIGNATURE _____
 204. PRINT NAME _____
 205. DATE _____
 206. SIGNATURE _____
 207. PRINT NAME _____
 208. DATE _____
 209. SIGNATURE _____
 210. PRINT NAME _____
 211. DATE _____
 212. SIGNATURE _____
 213. PRINT NAME _____
 214. DATE _____
 215. SIGNATURE _____
 216. PRINT NAME _____
 217. DATE _____
 218. SIGNATURE _____
 219. PRINT NAME _____
 220. DATE _____
 221. SIGNATURE _____
 222. PRINT NAME _____
 223. DATE _____
 224. SIGNATURE _____
 225. PRINT NAME _____
 226

[illegible]

[illegible][illegible]

FIGURE 51 (cont.)

1691 ----- 1740
 TCGTGCACCTGCGCTGAGTGTGTAGTGTGTCGAGCTGCTCAGGTCTTTCTTTTATGTCA
 AGGAAGTGCACGACTACTACACATGACGAGCTGCGAGTCCAGAAACAAAATACAGT
 a S C T G * * V C T S S S C S C L S F M S -
 b F A I A D E C V K R R A A Q V F L L C H -
 c L H W L M S V Y V V E L L R S F F Y V T -
 1741 ----- 1800
 CCGAGACCAAGTTTCAAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGAGCAAGT
 GCCTCTGGTGGCAAAAGTTTCTTCTCGAGCAAAAAGATGCGCTTCTCAGAGACCTGCTTCA
 a R R F R F K R T G S F S T G R V S G A S -
 b G D H V S K E Q A L F L P E E C L E Q V -
 c E T T F Q K N R L F F Y R K S V W S K L -
 1801 ----- 1860
 TCCAAACCATTTGGAATCAGACAGCACTTGAAGAGGGTGCAGCTGCGGGAGCTGTGCGAAG
 ACGTTTGTAACTTAGTCTGTCTGTAACCTTCTCCGAGGTGAGCGCTGAGAGCGCTTC
 a C K A L E S D S T * R G C S C G S C R K -
 b A K H W N Q T A L E E C A A A C A V G S -
 c Q S I G T R Q H L K R V Q L R E I S E A -
 1861 ----- 1920
 CAGAGTTCAGGCGAGCTGGGAAGCCAGGCGCGCGCTGCTGACGTCCAGACTCCGCTTCA
 GTCTCCACTCCGCTGTACCCCTTGTGTCGGGCGGGCAGGACTCCAGGTCTGACCCGAAAGT
 a Q R S G S I G K P G P P C * R F D S A S -
 b E G Q A A S C S Q A R P A D V Q T P L H -
 c E V R Q H R E A R P A L L T S R L R F I -
 1921 ----- 1980
 TCCCGAGCCTGACGGGCTGCGGCGGATTGTGAACATGACTAGCTGCTGGGAGCCAGAA
 AGGAGTTGGAGTGGCGAGCGCGGCTAACACTTGT/CCTGATGCAGCACCCCTGGGTCTT
 a S P S L T G C G R L * T W T T S W E P E -
 b P Q A * R A A A D C E H C L R R G S Q N
 c P K P D G L R P I V N M D Y V V G A K T
 1981 ----- 2040
 CGTTCCGACAGAAAGAGAGCGCCGAGCGTCTACCTCGAGGGTGAAGGCACTTTCAGAG
 GCAAGGCTCTCTTTTCTCCCGGCTGACAGAGTGGAGCTCCCACTTCCGAGACAAGTGGC
 a R S A E K R C P S V S P R G * R H C S A -
 b V P Q R K E G R A S H L E G E G T V Q R -
 c F R R E K R A E P L T S R V K A L F S V -
 2041 ----- 2100
 TGCTCACTACGAGCGCGCGCGCGCGCTCTCGGCGCGCTTGTGTGAGGCTGAG
 ACGAGTGTATCTGCGCGCGCGCGCGCGCGCGAGCACCGCGGAGACATGACCTCGAGC
 a C S T T S G R G A P A S W A P L C W A W -
 b A Q L R A G A A P R P P G R L C A G P G -
 c L N Y E R A R R P G L L G A S V L G L D -

1691
 1741
 1801
 1861
 1921
 1981
 2041

[illegible][illegible]

THE UNIVERSITY OF CHICAGO
 LIBRARY
 540 EAST 57TH STREET
 CHICAGO, ILL. 60637
 TEL: 773-936-5000
 FAX: 773-936-5000
 WWW: WWW.CHICAGO.EDU

[illegible]

10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----

	GTCGGAAAGCTCCCGGAGACGCCTACTGCCCTGGAGGCGCAGGCAACCCCGTCACTGC	
3361	-----+-----+	3420
A	CAGCCTTCGAGGSGCCCCGTGCTGCGACTGACCGGACCCTCCGGCCTCGGTTCCGCCCCTGACG	
B	V G S S R G R R * I F W P P Q F T R H C -	
C	S E A F G D D A D C P G G R S Q P G T A -	
D	R K L F G T T L T A L E A A A N P A L F -	
E	CCTCAGACATTCAGACCATCCCTGGACTGAATGCCACCTGCCACAGCCAGGCGAGAGCA	
3421	-----+-----+	3480
A	GGAGTCTGAAGTTCTGCTAGGACCTGACTACCGGTGGCGCGGTSTCGGTCGCCCTCTCTCT	
B	P Q T S R P S W T D G H P P T A E P R A -	
C	L R L Q D H P G L M A T R P Q P G R E Q -	
D	S D F K T I L D * W P F A H S Q A E S R -	
E	GACACCAGCAGGCCCTCTCAAGSCCGGGCTCTACGTCCAGGGAGGGAGGGCGGCCCACAC	
3481	-----+-----+	3540
A	CTGTGCTGCTGCGGACAGTGCCGGCCGAGATGCAGGGTCCCTCCCTCCCGCGCGGTGTG	
B	D T S S F V T P G S T S Q G G R G G P H -	
C	T F A A L S R R A L R P R E G G A A H T -	
D	H Q Q P C H A G L Y V P G R E G E P T P -	
E	CCAGGTCGGCACCCTCGGAGCTCTCAAGCCTGAGTGAGTGTTTTGGCCGAGGGCTGCATGT	
3541	-----+-----+	3600
A	GGTCCGGCGGTGGCGAACCTTCAGACTCCGACTCACTCACAAACGGGTCCGGACGTACA	
B	F G P H R W E S E A * V S V W P R P A C -	
C	Q A R T A G S L R P E * V P G R G L H V -	
D	K P A P L G V * G L S E C L A E A C M S -	
E	CCGGCTAAGCTGAGTGTCTCGGCTGAGGCCTAGCGAGTGTCCACCCAAGCGCTGAGTG	
3601	-----+-----+	3660
A	GGCGACTTCCGACTCACAGGCCGACTCCGGACTCGCTCACAGGTCCGTTCCCGACTCAC	
B	P A E G * V S G * G L S E C P A K C * V -	
C	R L F A E C P A E A * A S V Q P R A E C -	
D	G * R L S V R L R P E R V S S Q G L S V -	
E	TCCAGTACACCTTCGGCTCTTCACTTCCCCACARCTGCGCTGGCTCCACCCCCAGGSC	
3661	-----+-----+	3720
A	AGGTGTGTGTGACGGGACAGAAGTAACCGGTGTCCGACCGGAGCGGAGGTGCGATTCCGG	
B	S S T P A V F T S P Q A G A R L H P R A -	
C	P A H L P S S L P H R L A L G S T P G F -	
D	Q H T C E L H F P T G W R S A F P Q G Q -	
E	AGCTTTTCTCACCAAGAGCCCGGCTTCACTTCCCCACATAAGGATAGTCCATTCGCCAGA	
3721	-----+-----+	3780
A	TGGAAAAGGAGTGGTCTTCGGCCGAAGGTGAGGGGTGATCTTATCATCGCTAGCGGTCT	
B	S F S S P G A R L P L P T * E * S T P E	
C	A F P H Q E F G F H S P H R N S P S P D -	
D	L F L T R S P A S T P H I G I V H F Q T -	

FIGURE 51 (cont.)

3781 TTTGCGATTGTTTACCCGCTGCGCTGCGCTGCTTTTTCGCTTCGACCCCAACCATCCAGGTG 3840
AAGCGTATCAAGTGTCGGACCGGCACCGCAAGCAACCGAAGCGTGGCGGTGGTAGGTCGAC

```

      F A I V H F S P C P P L F S T P T I Q V -
      S F L F T P E P A L L C L P P P P S R W -
      F H C S P L A L P S F A F H P H H F G G -

```

GAGGACCTTCAGCAGGACCOCTGGAGACTCTTGGGAATTTGGAGTGACCAGGTGTGCCCTC
7841 +-----+ +-----+ +-----+ +-----+ +-----+ 3500
CTCTTAGACTCTTCCTGGGACCCCTGCAGACCOCTTAAGCTCACTGGTTTCACAAGCGGC

A T L E R T L G A L G I W S D Q R C A L -
L E P * E G P W E L W E F G V T K G V P C -
S D P E E D F G S S C N L E * P K V C L F V

2901 TACACAGGGGAGGACCCCTGCACCTGGATGGGGTCCCTGTGGGTCAAAATTGGGAGAGT 3960
ATGTGTCCCTCTCTGGACGTGGACCTACCCCCAGGGACACCTAGTTTAACTCTCTCTCA

A Y T G E D P A P G W G S L W V K L G G G -
 B T Q A R T L H L D G G P C G S N W G E V -
 C H R R C P C T W M C V P V G Q I G G R C

GCTGTGGGAGTAAATACTGAATATATGAGTTTTTCAGTTTTCAGAAAAAAGAAAAAGAAA
3961 -----+-----+-----+-----+-----, 4020
CCACACCCTCATTTTATCACTTATATACTCAGAAAAGCAGAAACTTTTITTTTTTTTTTTTT

```

a      A V G V K Y * I Y E F F S F E K K K K K
b      L W E * N T E Y M S F S V L K K K K K K
c      C G S K I L N I * V F Q P * K K K K K K

```

4021 AAAAAAAAAA
----- 4029
TTTTTTTTTT

а	к	к	к	-
б	к	к		-
с	к	к		-

FIGURE 52

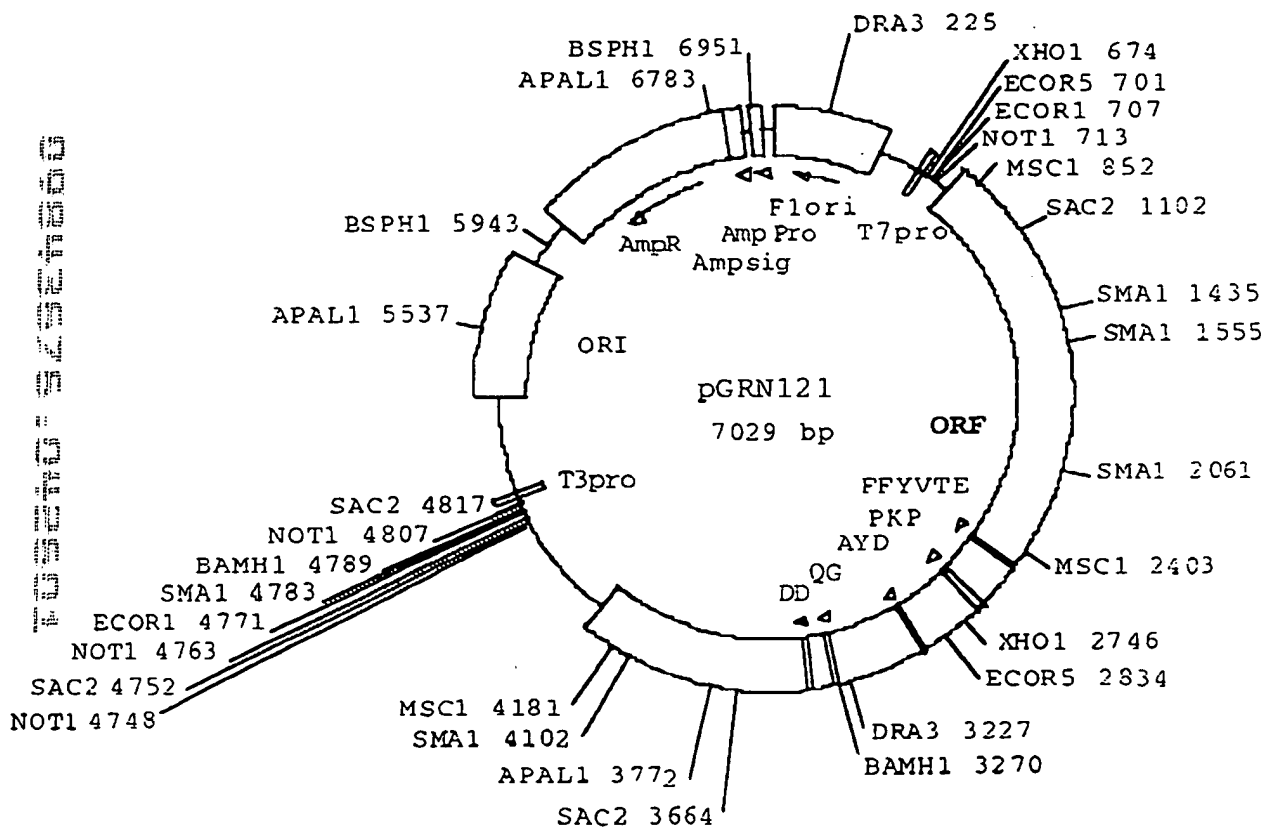


FIGURE 53

1
 met
 GCAGCGCTGCGTCCTGCTGCGCACGTGGGAAGCCCTGGCCCCGGCCACCCCCGCG ATG
 10
 pro arg ala pro arg cys arg ala val arg ser leu leu arg ser
 CCG CGC GCT CCC CGC TGC CGA GCC GTG CGC TCC CTG CTG CGC AGC
 20
 his tyr arg glu val leu pro leu ala thr phe val arg arg leu
 CAC TAC CGC GAG GTG CTG CCG CTG GCC ACG TTC GTG CGG CGC CTG
 30
 40
 gly pro gln gly trp arg leu val gln arg gly asp pro ala ala
 GGG CCC CAG GGC TGG CGG CTG GTG CAG CGC GGG GAC CCG GCG GCT
 50
 phe arg ala leu val ala gln cys leu val cys val pro trp asp
 TTC CGC GCG CTG GTG GCC CAG TGC CTG GTG TCC GTG CCC TGG GAC
 60
 70
 ala arg pro pro pro ala ala pro ser phe arg gln val ser cys
 GCA CGG CCG CCC CCC GCC GCC CCC TCC TTC CGC CAG GTG TCC TGC
 80
 leu lys glu leu val ala arg val leu gln arg leu cys glu arg
 CTG AAG GAG CTG GTG GCC CGA GTG CTG CAG AGG CTG TGC GAG CGC
 90
 100
 gly ala lys asn val leu ala phe gly phe ala leu leu asp gly
 GGC GCG AAG AAC GTG CTG GCC TTC GGC TTC GCG CTG CTG GAC GGG
 110
 ala arg gly gly pro pro glu ala phe thr thr ser val arg ser
 GCC CGC GGG GGC CCC CCC GAG GCC TTC ACC ACC AGC GTG CGC AGC
 120
 130
 tyr leu pro asn thr val thr asp ala leu arg gly ser gly ala
 TAC CTG CCC AAC ACG GTG ACC GAC GCA CTG CGG GGG AGC GGG GCG

[illegible]

610														
ser	glu	ala	glu	val	arg	gln	his	arg	glu	ala	arg	pro	ala	leu
TCG	GAA	GCA	GAG	GTC	AGG	CAG	CAT	CGG	GAA	GCC	AGG	CCC	GCC	CTG
620														
leu	thr	ser	arg	leu	arg	phe	ile	pro	lys	pro	asp	gly	leu	arg
CTG	ACG	TCC	AGA	CTC	CGC	TTC	ATC	CCC	AAG	CCT	GAC	GGG	CTG	CGG
630														
pro	ile	val	asn	met	asp	tyr	val	val	gly	ala	arg	thr	phe	arg
CCG	ATT	GTG	AAC	ATG	GAC	TAC	GTC	GTG	GGA	GCC	AGA	ACG	TTC	CGC
640														
arg	glu	lys	arg	ala	glu	arg	leu	thr	ser	arg	val	lys	ala	leu
AGA	GAA	AAG	AGG	GCC	GAG	CGT	CTC	ACC	TCG	AGG	GTG	AAG	GCA	CTG
650														
phe	ser	val	leu	asn	tyr	glu	arg	ala	arg	arg	pro	gly	leu	leu
TTC	AGC	GTG	CTC	AAC	TAC	GAG	CGG	GCG	CGG	CGC	CCC	GGC	CTC	CTG
660														
gly	ala	ser	val	leu	gly	leu	asp	asp	ile	his	arg	ala	trp	arg
GGC	GCC	TCT	GTG	CTG	GGC	CTG	GAC	GAT	ATC	CAC	AGG	GCC	TGG	CGC
670														
thr	phe	val	leu	arg	val	arg	ala	gln	asp	pro	pro	pro	glu	leu
ACC	TTC	GTG	CTG	CGT	GTG	CGG	GCC	CAG	GAC	CCG	CCG	CCT	GAG	CTG
680														
tyr	phe	val	lys	val	asp	val	thr	gly	ala	tyr	asp	thr	ile	pro
TAC	TTT	GTC	AAG	GTG	GAT	GTG	ACG	GGC	GCG	TAC	GAC	ACC	ATC	CCC
690														
gln	asp	arg	leu	thr	glu	val	ile	ala	ser	ile	ile	lys	pro	gln
CAG	GAC	AGG	CTC	ACG	GAG	GTC	ATC	GCC	AGC	ATC	ATC	AAA	CCC	CAG
700														
asn	thr	tyr	cys	val	arg	arg	tyr	ala	val	val	gln	lys	ala	ala
AAC	ACG	TAC	TGC	GTG	CGT	CGG	TAT	GCC	GTG	GTC	CAG	AAG	GCC	GCC
710														
his	gly	his	val	arg	lys	ala	phe	lys	ser	his	val	ser	thr	leu
CAT	GGG	CAC	GTC	CGC	AAG	GCC	TTC	AAG	AGC	CAC	GTC	TCT	ACC	TTG
720														
730														
740														
750														
760														

[illegible]

770														780	
thr	asp	leu	gln	pro	tyr	met	arg	gln	phe	val	ala	his	leu	gln	
ACA	GAC	CTC	CAG	CCG	TAC	ATG	CGA	CAG	TTC	GTG	GCT	CAC	CTG	CAG	
790															
glu	thr	ser	pro	leu	arg	asp	ala	val	val	ile	glu	gln	ser	ser	
GAG	ACC	AGC	CCG	CTG	AGG	GAT	GCC	GTC	GTC	ATC	GAG	CAG	AGC	TCC	
800														810	
ser	leu	asn	glu	ala	ser	ser	gly	leu	phe	asp	val	phe	leu	arg	
TCC	CTG	AAT	GAG	GCC	AGC	AGT	GGC	CTC	TTC	GAC	GTC	TTC	CTA	CGC	
820															
phe	met	cys	his	his	ala	val	arg	ile	arg	gly	lys	ser	tyr	val	
TTC	ATG	TGC	CAC	CAC	GCC	GTG	CGC	ATC	AGG	GGC	AAG	TCC	TAC	GTC	
830														840	
gln	cys	gln	gly	ile	pro	gln	gly	ser	ile	leu	ser	thr	leu	leu	
CAG	TGC	CAG	GGG	ATC	CCG	CAG	GGC	TCC	ATC	CTC	TCC	ACG	CTG	CTC	
850															
cys	ser	leu	cys	tyr	gly	asp	met	glu	asn	lys	leu	phe	ala	gly	
TGC	AGC	CTG	TGC	TAC	GGC	GAC	ATG	GAG	AAC	AAG	CTG	TTT	GCG	GGG	
860														870	
ile	arg	arg	asp	gly	leu	leu	leu	arg	leu	val	asp	asp	phe	leu	
ATT	CGG	CGG	GAC	GGG	CTG	CTC	CTG	CGT	TTG	GTG	GAT	GAT	TTC	TTG	
880															
leu	val	thr	pro	his	leu	thr	his	ala	lys	thr	phe	leu	arg	thr	
TTG	GTG	ACA	CCT	CAC	CTC	ACC	CAC	GCG	AAA	ACC	TTC	CTC	AGG	ACC	
890														900	
leu	val	arg	gly	val	pro	glu	tyr	gly	cys	val	val	asn	leu	arg	
CTG	GTC	CGA	GGT	GTC	CCT	GAG	TAT	GGC	TGC	GTG	GTG	AAC	TTG	CGG	
910															
lys	thr	val	val	asn	phe	pro	val	glu	asp	glu	ala	leu	gly	gly	
AAG	ACA	GTG	GTG	AAC	TTC	CCT	GTA	GAA	GAC	GAG	GCC	CTG	GGT	GGC	
920														930	
thr	ala	phe	val	gln	met	pro	ala	his	gly	leu	phe	pro	trp	cys	
ACG	GCT	TTT	GTT	CAG	ATG	CCG	GCC	CAC	GGC	CTA	TTC	CCC	TGG	TGC	

FIGURE 53 (cont.)

940
gly leu leu leu asp thr arg thr leu glu val gln ser asp tyr
GGC CTC CTG CTG GAT ACC CGG ACC CTG GAG GTG CAG AGC GAC TAC

950
ser ser tyr ala arg thr ser ile arg ala ser val thr phe asn
TCC AGC TAT GCC CGG ACC TCC ATC AGA GCC AGT GTC ACC TTC AAC

960
arg gly phe lys ala gly arg asn met arg arg lys leu phe gly
CGC GGC TTC AAG GCT GGG AGG AAC ATG CGT CGC AAA CTC TTT GGG

970
val leu arg leu lys cys his ser leu phe leu asp leu gln val
GTC TTG CGG CTG AAG TGT CAC AGC CTG TTT CTG GAT TTG CAG GTG

980
asn ser leu gln thr val cys thr asn ile tyr lys ile leu leu
AAC AGC CTC CAG ACG GTG TGC ACC AAC ATC TAC AAG ATC CTC CTG

990
leu gln ala tyr arg phe his ala cys val leu gln leu pro phe
CTG CAG GCG TAC AGG TTT CAC GCA TGT GTG CTG CAG CTC CCA TTT

1000
his gln gln val trp lys asn pro thr phe phe leu arg val ile
CAT CAG CAA GTT TGG AAG AAC CCC ACA TTT TTC CTG CGC GTC ATC

1010
ser asp thr ala ser leu cys tyr ser ile leu lys ala lys asn
TCT GAC ACG GCC TCC CTC TGC TAC TCC ATC CTG AAA GCC AAG AAC

1020
ala gly met ser leu gly ala lys gly ala ala gly pro leu pro
GCA GGG ATG TCG CTG GGG GCC AAG GGC GCC GCC GGC CCT CTG CCC

1030
ser glu ala val gln trp leu cys his gln ala phe leu leu lys
TCC GAG GCC GTG CAG TGG CTG TGC CAC CAA GCA TTC CTG CTC AAG

1040
leu thr arg his arg val thr tyr val pro leu leu gly ser leu
CTG ACT CGA CAC CGT GTC ACC TAC GTG CCA CTC CTG GGG TCA CTC

1050
1060
1070
1080
1090

FIGURE 53 (cont.)

1100 1110
 arg thr ala gln thr gln leu ser arg lys leu pro gly thr thr
 AGG ACA GCC CAG ACG CAG CTG AGT CGG AAG CTC CCG GGG ACG ACG
 1120
 leu thr ala leu glu ala ala ala asn pro ala leu pro ser asp
 CTG ACT GCC CTG GAG GCC GCA GCC AAC CCG GCA CTG CCC TCA GAC
 1130 1132
 phe lys thr ile leu asp OP
 TTC AAG ACC ATC CTG GAC TGA TGGCCACCCGCCACAGCCAGGCCGAGAGCAGA
 CACCAGCAGCCCCTGTCACGCCGGGCTCTACGTCCCAGGGAGGGAGGGGCGGCCACACCC
 AGGCCCCGCACCGCTGGGAGTCTGAGGCCTGAGTGAGTGTTTGGCCGAGGCCTGCATGTCC
 GGCTGAAGGCTGAGTGTCCGGCTGAGGCCTGAGCGAGTGTCCAGCCAAGGGCTGAGTGTC
 CAGCACACCTGCCGTCTTCACTTCCCCACAGGCTGGCGCTCGGCTCCACCCCAGGGCCAG
 CTTTTCYTCACCAGGAGCCCGGCTTCCACTCCCCACATAGGAATAGTCCATCCCCAGATT
 CGCCATTGTTACCCCYTCGCCCTGCCYTCCTTTGCCTTCCACCCCCACCATCCAGGTGGA
 GACCCTGAGAAGGACCCTGGGAGCTCTGGGAATTTGGAGTGACCAAAGGTGTGCCCTGTA
 CACAGGCGAGGACCCTGCACCTGGATGGGGGTCCCTGTGGGTCAAATTGGGGGGAGGTGC
 TGTGGGAGTAAAATACTGAATATATGAGTTTTTCAGTTTTGRAAAAAAAAAAAAAAAAAAA
 AAAAAAAAAA

FIGURE 54

KPN1 17050
KPN1 18560
ECOR1 19475
SAC1 19652
SAC1 21083
ECOR1 23539
SAC1 24607
SAC1 29043

<-TCP